NTNU
Innovation and Creativity

Automatic diagnosis of ultrasound images using standard view planes of fetal anatomy

Jan Ødegård
Anders Østen

Master of Science in Computer Science
Submission date: June 2006
Supervisor: Ketil Bø, IDI

Norwegian University of Science and Technology
Department of Computer and Information Science

Problem Description

Develop a modular system for automatic diagnosis of the fetus using ultrasound images (standard view planes) obtained at the routine ultrasonography at 18 weeks. The system shall perform extraction and recognition of the features from the standard view planes. Based on the combination of this knowledge and medical domain knowledge, a diagnosis of the fetus can be automatically made. The diagnosis is to be presented to the user performing the examination.

Assignment given: 20. January 2006
Supervisor: Ketil Bø, IDI

ABSTRACT

The use of ultrasound has revolutionised the area of clinical fetal examinations. The possibility of detecting congenital abnormalities at an early stage of the pregnancy is highly important to maximise the chances of correcting the defect before it becomes life-threatening. The problems related to the routine procedure is its complexity and the fact that it requires a lot of knowledge about fetal anatomy. Because of the lack of training among midwives, especially in less developed countries, the results of the examinations are often limited. In addition, the quality of the ultrasound equipment is often restricted. These limitations imply the need for a standardised procedure for the examination to decrease the amount of time required, as well as an automatic method for proposing the diagnosis of the fetus.

This thesis has proposed a solution for automatically making a diagnosis based on the contents of extracted ultrasound images. Based on the concept of standard view planes, a list of predefined images are obtained of the fetus during the routine ultrasonography. These images contain the most important organs to examine, and most common congenital abnormalities are therefore detectable in this set. In order to perform the analysis of the images, medical domain knowledge must be obtained and stored to enable reasoning about the findings in the ultrasound images. The findings are extracted through segmentation and each object is given a unique description. An organ database is developed to store descriptions about existing organs to recognise the extracted objects. Once the organs have been identified, a CBR system is applied to analyse the total contents of one standard view plane. The CBR system uses domain knowledge from the medical domain as well as previously solved problems to identify possible abnormalities in the case describing the standard view plane. When a solution is obtained, it is stored for later retrieval. This causes the reliability of future examinations to increase, because of the constant expansion of the knowledge base.

The foundation of standard view planes ensures an effective procedure and the amount of training needed to learn the procedure is minimised due to the automatic extraction and analysis of the contents of the standard view plane. The midwife only has to learn which standard view planes to obtain, not the analysis of their contents.

Preface

This master thesis documents the work performed by Jan Ødegård and Anders Østen in the period from January 20th to June 16th, 2006. The main goal has been to establish a system for obtaining a defined set of ultrasound images during the routine ultrasonography at 18 weeks of pregnancy, and to automatically make a diagnosis based on their contents. The work has been performed at the Norwegian University of Science and Technology (NTNU), the Department of Computer and Information Science (IDI). The work is an extension of the in-depth study performed during the autumn semester of 2005.

We would like to thank our supervisor Ketil Bø for his guidance and useful discussions during our work on this thesis.

Trondheim, June 16, 2006.

Jan Ødegård

Anders Østen

Contents

LIST OF FIGURES

List of Tables

List of Listings

Chapter 1 Introduction

This chapter will provide the background and motivation for this thesis, accompanied by a brief description of the task at hand. An overview of the remaining thesis is provided by the last section of this chapter.

1.1 Background and motivation

The use of ultrasound has revolutionised the area of clinical fetal examinations. This technology makes it possible to detect defects and abnormalities at an early stage of the pregnancy, increasing the chances of a successful treatment. At 18 weeks of pregnancy, most characteristics of the fetus are fully developed, making it a suitable time to perform the examination.

The main problem of the examination is its complexity. It requires a lot of skill and training to be accurately executed, which is a problem especially in less developed countries. The economy and knowledge to support the education of midwifes are often missing, allowing easily detectable abnormalities to become life-threatening. Assistance from intelligent systems analysing the ultrasound images during the examination would be of great help and support, enabling the midwife to focus on learning the procedures of the examination.

This thesis proposes a possible solution to how automatic analysis of ultrasound images can be accomplished. The concept of standard view planes of fetal anatomy, as defined in [1], and intelligent reasoning using general domain knowledge are utilised to enable this analysis.

1.2 Task definition

This section will provide a description of the main task of this thesis. This task is decomposed into three subtasks for easier application and modularity.

The main task of this thesis is to utilise knowledge about fetal anatomy and fetal abnormalities to aid the routine fetal ultrasound examination at 18 weeks of pregnancy. The anatomy and abnormalities will be described using a reference system of standard view planes of fetal anatomy. By incorporating intelligent reasoning, the general knowledge available in the medical domain is utilised. To examine the feasibility of the solution, a prototype implementation will be developed.

To be able to establish a reference system of standard view planes, a thorough description of each view plane is required. The first subtask will be to develop such descriptions for the standard view planes applicable to this thesis. The descriptions should contain information about the anatomical content and the normalities and abnormalities commonly observed in the view planes.

The establishment of a reference system of standard view planes, makes it possible to create associations between the information contained in ultrasound images and the knowledge needed to perform an automatic analysis. The second subtask is the extraction of this information as features from the standard view planes. Because ultrasound images are inherently noisy, several steps of processing are required in order to discover the true features of an image. The features obtained must be described in a way suitable for knowledge assessment.

The final subtask requiring elaboration is the actual analysis of the features extracted from the images. The knowledge gained during the examination must be analysed using the stored knowledge about the fetal anatomy and abnormalities. Based on this analysis, a diagnosis can be made and presented to a medical expert for further evaluation. This ensures that the user performing the ultrasound examination is only responsible for executing the examination, without having to consider the difficult process of analysing the images obtained.

1.3 Overview

This section presents an overview of this thesis with a brief description of each chapter.

Chapter 2 will present the standard view planes of choice accompanied by reasons for choosing them. Each standard view plane has a feature description, which consists of anatomical landmarks and anatomical descriptors, and a medical basis which discusses the importance of the described features. The most common abnormalities recognised in each of the chosen standard view planes are described along with their symptoms.

Chapter 3 will elaborate the task of extracting the features described in chapter 2. Different methods of segmentation are presented, and an appropriate segmentation method is chosen for the problem at hand. Ultrasound images are inherently noisy, which forces the solution to consider this noise when identifying the features in the images. This chapter will present noise characteristics and possible solutions to this problem. The extracted features require unique descriptions before being submitted to a procedure of matching features and organs. Different approaches for creating such descriptions will be discussed, and a suitable one is chosen. Finally, a matching procedure is proposed to understand the extracted descriptions.

After completing extraction and understanding of the features, chapter 4 discusses different techniques applicable for image analysis. This analysis is based on the gathering and reasoning of the knowledge extracted from the standard view planes and the knowledge inherently contained in the medical domain. Two methods for image analysis are described, and the one satisfying the constraints of the problem domain is chosen.

Chapter 5 summarises chapters 2-4 and defines a list of components required to perform automatic analysis of ultrasound images. The practical details of each component is explained, and examples from the developed prototype are presented. The workflow between the components is also presented in this chapter.

An evaluation of the work done in this thesis is given in chapter 6, and future work is presented in chapter 7.

A more detailed explanation of the chosen segmentation method is given in appendix A.

Appendix B contains the mathematical foundation for the descriptors used to represent the medical regions.

Chapter 2
Standard view planes and features

This chapter will further elaborate on the first subtask presented in chapter 1, the development of a thorough description of the fetal anatomy through the use of standard view planes. First, a short outline of the concept of standard view planes is given in section 2.1. This thesis will use only a small subset of the standard view planes previously presented in [1], as a complete description of all standard view planes is beyond the scope. A small subset ensures that focus is maintained on the analysis of the image contents.

The standard view planes of choice are the biparietal diameter (BPD) (section 2.2), abdominal diameter (AD) (section 2.3), and the four chamber view (4CH) (section 2.4). Reasons for choosing these specific standard view planes are elaborated in their corresponding sections.

The anatomical content of each standard view plane will be described using anatomical landmarks and descriptors. Each section will contain a medical basis for including the particular landmarks, which are often closely related to abnormalities present in the standard view plane. The abnormalities presented in each section are further described in section 2.5.

2.1 Standard view planes of fetal anatomy

This section will shortly outline the concept of standard view planes. The main reason for their suitability when working with image analysis is also explained.

A standard view plane of fetal anatomy is, as stated in [1], an ultrasound image which is defined by its contents rather than its pixels. The contents is a set of features that should be present in the standard view plane. A definition based on these features instead of for instance transducer placement, ensures that the standard view plane is independent of the movement, size, and orientation of the fetus. A defined set of standard view planes applicable during the routine ultrasound examination ensures a standardised procedure which is easy to initiate in a new environment.

In order to perform intelligent image analysis, the knowledge obtained during the ultrasound examination must be stored and compared to already existing domain knowledge. By extracting the domain knowledge based on standard view planes, as well as storing the knowledge obtained during the routine examination (through the use of standard view planes), a common representation form of the knowledge has been established. This eases the task of the analysis tool, as both the unknown and the already defined knowledge are categorised in a similar way.

2.2 Biparietal diameter (BPD)

The biparietal diameter measures, roughly spoken, the size (diameter) of the fetal skull. It is found by obtaining a section through the parietal eminences, perpendicular to falx cerebri [2]. This view plane is often referred to as the transthalamic plane [3]. Various reference points on the skull are used for the measurement (inner-outer, outer-outer, middle-to-middle), depending on the type of transducer applied. It appears that the large variation of normal measurements at any gestational age makes the method of measurement less important.

The BPD is an important measurement, which is subsequently used for many purposes. The main purpose is the estimation of the gestational age of the fetus, which indicates whether the fetus develops as expected or not. There is a strong correlation between the age and the development of certain organs of the fetus. If these organs are not present, or not fully developed at the right age, it should be possible to make a diagnosis, and hopefully carry out a correcting surgical intervention. The BPD should be measured before the 25th week of pregnancy, as the variations in the size of the fetal head increase after this stage. In the late second and third trimester the BPD is merely used in weight-prediction equations. An example of a correctly obtained BPD standard view plane is shown in figure 2.1.

Figure 2.1: Illustration of the BPD standard view plane. Courtesy of [4].

2.2.1 Basis of choice

The BPD standard view plane has been chosen for several important reasons:

- Historical importance and recognition of the BPD measurement
- Suitability of discovery
- Integrated part of usual examination
- Landmark importance when assessing the gestational growth

Since the start of cephalometry in the early 1970s, the BPD measurement has been very important to assess the rate of fetal growth. Its close relation to the discovery of clinically disabling conditions at an early stage makes it a substantial part of every ultrasonographic examination. Through medical development and research the BPD has become an established part of modern fetal analysis.

Another reason for choosing the BPD standard view plane is its well documented routines for discovery. It is oriented according to the midline echo of the fetal head, which represents the medial aspect of each cerebral hemisphere. This midline echo will bisect the standard view plane on its longest axis when it is obtained transversely. Because of this precise location of the standard view plane, high reproducibility is ensured.

As mentioned above, the BPD measurement is an integrated part of the routine examination of pregnant women in Norway. The BPD standard view plane is defined at the exact same location as where the measurement is taken and is easily obtained for further processing. It is also possible to imagine an automatic measurement of the BPD whenever the standard view plane is obtained.

The anatomical landmarks found in the BPD standard view plane are very important indicators on healthy fetal growth. A lot of abnormalities connected to the head and spine will have symptoms visible in organs or landmarks present in the BPD standard view plane.

2.2.2 Feature description

This section will describe the content of the BPD standard view plane and its importance. The content will be partitioned into anatomical landmarks and anatomical descriptors. The anatomical landmarks describe the medical features of the image, i.e., particular organs. Anatomical descriptors describe the quantitative measurements that can be obtained through the image processing pipeline. The next sections will define the anatomical landmarks and the descriptors of the BPD standard view plane.

2.2.2.1 Anatomical landmarks

The most important content of any standard view plane is its anatomical landmarks. Anatomical landmarks are defined as organs that must always be present in the standard view plane.

The most important landmarks of the BPD standard view plane are:

- Midline Falx
- Cavum Septum Pellucidum
- Thalami
- Lateral cerebral ventricles

The midline falx, often referred to as the falx cerebri, is a dense band of fibrous tissue that separate the frontal cerebral lobes. It can be visualised as a short midline echo between the thalami and the sinciput.

The thalami are located in the center of the brain, one beneath each cerebral hemisphere, next to the third ventricle. The operational task of the thalami is to act as relay stations, receiving sensory input from the body or other parts of the brain and routing the proper signals to the cerebral cortex. The thalami should be located symmetrically on each side of the midline falx.

The septum pellucidum is a thin partition separating the lateral ventricles as the corpus callosum is growing. If the walls of this partition are not fused, a fluid-filled space called a cavum will emerge. The walls of the cavum septum pellucidum are parallel and the distance between them should be about 1 cm. The cavum is an enclosed space and not part of the ventricular system of the brain, nor is it connected to the subarachnoid space. During the routine ultrasonography it should be reported whether the walls of the cavum septum pellucidum are bowing or not. Bowing walls might indicate a cyst in the cavum. When measuring the occipitofrontal distance of the BPD standard view plane (see section 2.2.2.2), the cavum septum pellucidum is located one third into the brain along the line used for this measurement.

The lateral cerebral ventricles should be located in the cerebral hemispheres, extending into the occipital and temporal horns. They are part of the ventricular fluid conducting system and connected to the third ventricle via the interventricular foramina. Each lateral ventricle consists of several parts, as shown in figure 2.2. The most important features to notice are the anterior and posterior horns, which are located respectively in the frontal and occipital lobes of the brain. The anterior horns of the lateral ventricle should be separated by the triangular membranes of the septum pellucidum.

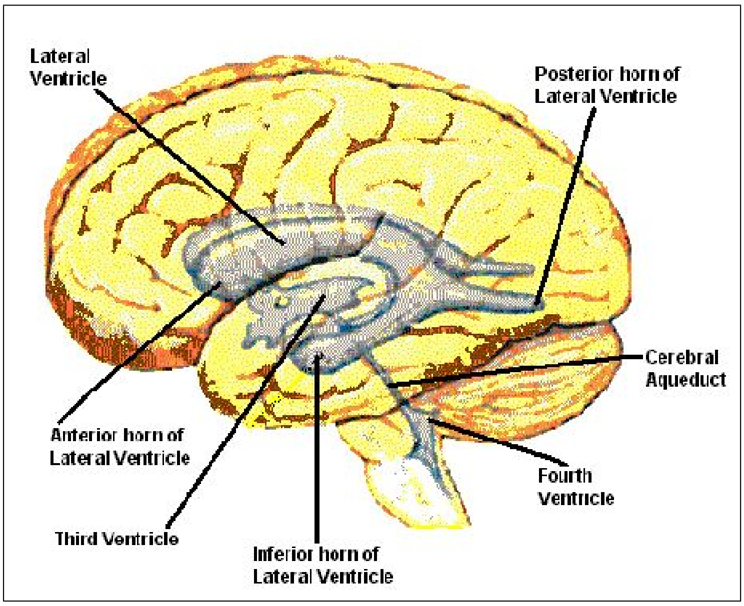


Figure 2.2: The ventricular system of the brain. Courtesy of [5].

2.2.2.2 Anatomical descriptors

Anatomical descriptors are quantitative measurements and descriptions that can be obtained by means of image processing and image enhancement. Considering the BPD standard view plane, the anatomical descriptors (displayed in figure 2.3) are:

- The biparietal diameter (BPD)
- The occipitofrontal diameter (OFD)

- Anterior cerebral ventricle diameter (Va)
- Posterior cerebral ventricle diameter (Vp)
- Hemispheric width (Hem)
- Elliptic shape

Figure 2.3: Illustration of the anatomical descriptors (BPD, OFD, Hem, Va and Vp) of the BPD standard view plane. Courtesy of [6].

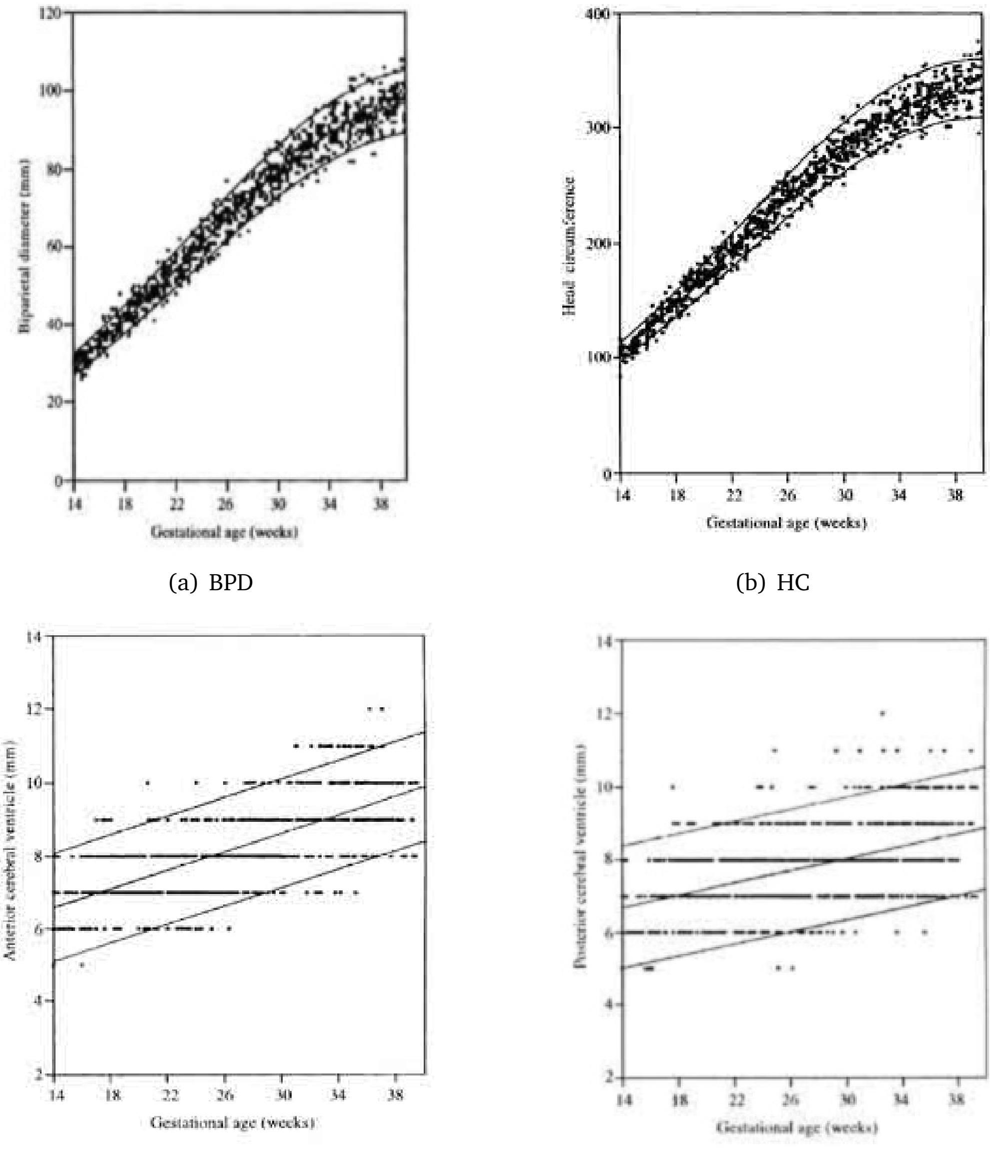
The BPD measurement is the main asset and very foundation of this standard view plane and thus the most important descriptor. At a gestational age of 18 weeks, the BPD should in most cases be between 35 mm and 47 mm. A nomogram describing the relationship between BPD and gestational age is shown in figure 2.4(a).

The OFD should be measured from the mid-frontal to the mid-occipital bone to minimize the inclusion of artifacts. Its measurement should be perpendicular to the BPD. The BPD and OFD can be combined into a third measurement, the head circumference (HC). The equation for calculating the HC is $HC = (BPD + OFD) * \frac{\pi}{2}$. A nomogram describing the relationship between the head circumference and gestational age is shown in figure 2.4(b).

The anterior (Va) and posterior (Vp) cerebral ventricle diameter are measured respectively as the distance between the wall of the anterior horn and the midline and the distance between the medial and lateral walls of the posterior horn. At a gestational age of 18 weeks, the Va should be between 5.7 mm and 8.6 mm and Vp should be between 5.4 mm and 8.8 mm. More details on their values can be found in the nomograms presented in figure 2.4(c) and 2.4(d).

The hemispheric distance should be measured from the midline to the inner border of the skull. This descriptor is not used as a separate measurement, but rather in comparison with Va and Vp. At 18 weeks of gestational age, the ratio Va/Hem should be in the range of 0.29 to 0.41 and the ratio Vp/Hem should be in the range of 0.27 to 0.42. These ratios are illustrated in figures 2.5(a) and 2.5(b).

Possibly the most obvious descriptor of the BPD standard view plane is its elliptic shape. Its size will vary across the population, but some kind of elliptic shape should always be present in a healthy fetus.



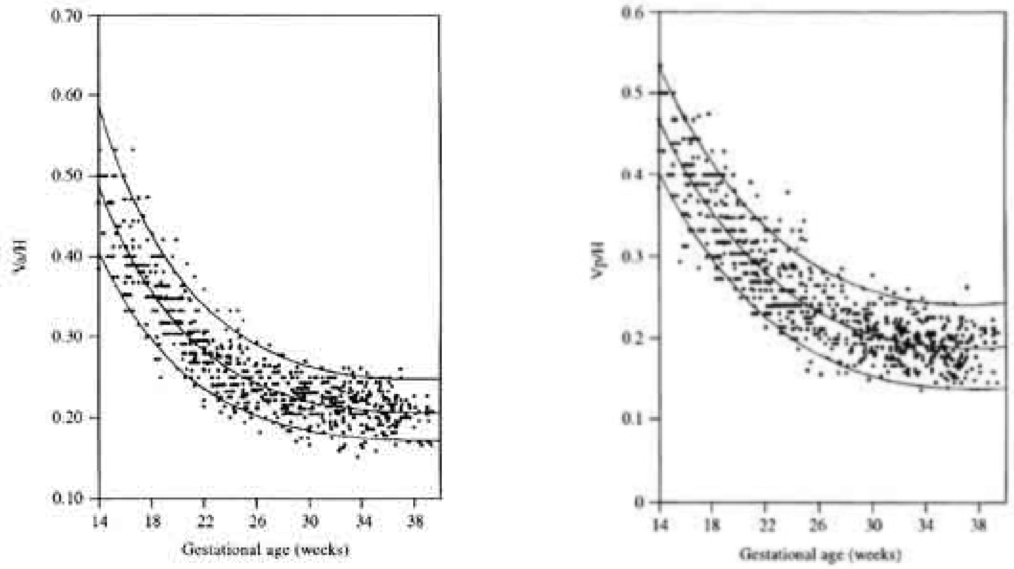
(a) BPD

(b) HC

(c) Va

(d) Vp

Figure 2.4: Nomograms of anatomical descriptors (I). Courtesy of [7].



(a) Va/H ratio

(b) Vp/H ratio

Figure 2.5: Nomograms of anatomical descriptors (II). Courtesy of [7].

2.2.2.3 Medical basis

This section will discuss the importance of including and identifying each of the anatomical landmarks mentioned in section 2.2.2.1. The abnormalities indicated by missing landmarks are also discussed. This section presents some of the symptoms of abnormalities found in the BPD standard view plane, while a more comprehensive list of symptoms is given in section 2.5.1.

As mentioned in section 2.2.2.1, the midline falx separates the cerebral hemispheres and should always be found in healthy fetuses. To avoid misinterpretations, a thorough scan should be performed before conclusions about its presence are drawn. A missing midline falx or echo indicates that the cerebral hemispheres are either completely connected or less connected than expected. The most common example of a missing or abnormal midline falx is holoprosencephaly (2.5.1.1). This is a complex malformation which concerns an abnormal cleavage of the cerebral hemispheres. The abnormality is categorised as either alobar, semi lobar, or lobar. The most severe category is the alobar, in which there is no cleavage between the hemispheres, resulting in a monoventricular cavity and a missing midline falx.

Holoprosencephaly also affects the thalami the same way it affects the midline falx. The alobar type results in a fusion of the thalami and the semi lobar type results in a partial fusion of the two.

The presence of the cavum septum pellucidum should always be confirmed. A missing septum pellucidum can indicate several abnormalities. Hydranencephaly (2.5.1.2) is a severe brain abnormality where large parts of the brain structures above the brain stem are absent and only cerebrospinal fluid is contained in the cranium. A central linear structure with the remains of the brain might be present, but often both the septum pellucidum and the midline falx are completely absent [8].

The agenesis of the corpus callosum (ACC) (2.5.1.5) is also an abnormality possibly indicated by a missing septum pellucidum. ACC is failure of the callosal fibers to form the corpus callosum in the midline between the cerebral hemispheres. The septum pellucidum has a close anatomical relationship with the corpus callosum and will also be absent in most cases. There is no evidence that the septum pellucidum cannot develop independently without the corpus callosum, but a usual claim is that there is a dependency between them. The diagnosis is further strengthened if the lateral ventricles are found to be largely separated.

If the skull is found absent when examining the BPD standard view plane, anencephaly (2.5.1.3) must be considered. Because of a defective closure of the neural tube, the forebrain, skull, and surrounding skin are more or less missing.

Spinal dysraphism (2.5.1.4) is an abnormality that does not originate in the BPD standard view plane, but symptoms are nevertheless found. The main symptom found when analysing the BPD standard view plane is the distinct lemon, or bullet shaped head. The special shape is most easily found slightly superior to where the BPD is measured, but one should still be able to suspect it when examining this standard view plane. In addition to the irregularly shaped head, the ventricles may also be dilated. If the head is both irregularly shaped, and there is dilation within one or more of the ventricles, one should immediately suspect spinal dysraphism.

The lateral ventricles are, as mentioned in section 2.2.2.1, part of the ventricular fluid conducting system of the brain. They are also the largest of the ventricles. The main reason for including them as one of the anatomical landmarks is that a lot of abnormalities have symptoms contained in, or in relation to, these structures. If the ventricles are found to be dilated, the most commonly related abnormality would be hydrocephalus (2.5.1.6). Hydrocephalus is a condition in which there is an excessive accumulation of fluid in the fetal brain causing the lateral ventricles to dilate. When the dilation is 10-15 mm the condition is called ventriculomegaly. Ventriculomegaly is just an increased size of the ventricles, without any increased accumulation of cerebrospinal fluid. It can, however, develop into hydrocephalus at a later stage. Between 15 and 40 weeks of gestational age the normal size of the lateral ventricles should be no more than 10 mm.

2.3 Mean abdominal diameter (MAD)

The mean abdominal diameter measures, as the name implies, the diameter of the abdomen. It is obtained with a transverse section through the abdomen at the height of the umbilical vein [2]. Because the abdomen is an organ which is easily deformable, two different measurements are needed to correctly calculate its diameter. The abdominal transversal diameter (ATD), and the abdominal anteroposterior diameter (APD) [2] are used to calculate the mean abdominal diameter: $MAD = \frac{1}{2}(APD + ATD)$. These are angled perpendicularly on each other, and together they provide a more robust estimate of the gestational age than BPD. The main problem when obtaining this measurement is the fact that both the ATD and APD are needed to produce a result. This implies that the user needs more knowledge and the chances of making a mistake is larger than using for instance BPD, which depends on only one measurement. A trade-off between BPD and MAD is needed. BPD is less prone to measuring errors, while MAD is more accurate, but at the same time includes more uncertainty concerning the measurement itself.

The reason for choosing mean abdominal diameter instead of abdominal circumference (AC) is mainly the recommendation by Eik-Nes in [2]. He states that the diameter is easier to measure compared to the circumference. Although measuring the circumference gives a more accurate measurement of the gestational age, the additional complications do not justify it.

2.3.1 Basis of choice

The MAD standard view plane has been chosen for several important reasons:

- Integrated part of the examination
- Well-defined anatomical landmarks
- Reliable quantitative measurements
- Independence of fetal orientation

The mean abdominal diameter is, as the BPD, measured during every routine examination in Norway. This means that the MAD standard view plane is defined at the same location as the measurement, implying no requirement of further training. Being a part of numerous examination guidelines, the MAD standard view plane should be easily integrated within the existing procedures[1].

Because of the comprehensive use of this view plane, it is well-defined in the medical literature. The landmarks needed to accurately define the complete structure are known. This eases the task of choosing the features needed.

Another reason for choosing MAD as one of the standard view planes of this thesis, is because it contains known quantitative measurements. These measurements are defined and given in nomograms stating the diameter at certain gestational ages. This means that the diameter can be immediately verified when performing the routine examination.

As mentioned earlier, the MAD is an integrated part of the Norwegian routine ultrasound examination. It requires a lot a training to visualise with a transducer, but it can always be found no matter how the fetus is situated. Provided with the correct training, the image is always obtainable at the routine examination.

2.3.2 Feature description

This section will contain a description of the features needed to define the standard view plane. These features will be divided into the same two categories as the BPD standard view plane; anatomical landmarks and anatomical descriptors. Each of the next sections define these features.

2.3.2.1 Anatomical landmarks

To precisely define each standard view plane, a set of anatomical landmarks must be defined. The main purpose of this set is to state which objects must be present in the standard view

[1]Abdominal circumference is sometimes used instead of MAD, but the standard view plane is almost the same.

plane. In addition, the inclusion of a thorough description of each landmark is used to determine whether or not the particular organ, described by the anatomical landmark, is healthy. If some are missing or deformed it could be a sign of an abnormality. It is, however, important to make sure that the view plane is obtained at the correct location. If the transducer is slightly out of position, a landmark might disappear in the image, although the fetus is perfectly healthy.

The most important anatomical landmarks of the MAD standard view plane are:

- Umbilical vein
- True transverse of spine
- Symmetrical ribs
- Stomach bubble

The umbilical vein should be angled perpendicular to the view plane. If an oblong section of the vein is visualised, the view plane is oblique, and the measurement will not be accurate. In addition, the vein should be positioned one third of the distance between the abdominal walls, on a line crossing the vein and the spine. This will help ensure a correctly obtained standard view plane.

A true transverse of the spine should be present on the opposite side of the umbilical vein. The three ossification centres of the vertebra should be visible to make sure that the view plane is taking at a perfectly transverse section through the fetus. If the image is obtained without fulfilling this criterion, the MAD measurement, in particular, will be unreliable.

The third landmark to look for is the appearance of symmetrical ribs. If the ribs are not correctly aligned, the transducer must be angled sideways in order to obtain a pair of symmetrically aligned ribs [9].

The fourth and final landmark included is the stomach bubble. It is visualised as a dark circular structure. The abnormalities related to the stomach bubble are further elaborated in section 2.3.2.3.

An illustration of a correctly obtained MAD standard view plane with all the necessary landmarks can be seen in figure 2.6.

2.3.2.2 Anatomical descriptors

As mentioned in section 2.2.2.2, the anatomical descriptors will be extracted during the image enhancement and image processing stages. In the MAD standard view plane, there are mainly three descriptors that are used.

- The mean abdominal diameter (MAD)
  - Abdominal transversal diameter (ATD)
  - Abdominal anteroposterior diameter (APD)
- Circular abdomen

Figure 2.6: Illustration of the MAD standard view plane.

- The umbilical vein should be positioned one third of the distance between the abdominal walls, along a line crossing the spine

The most important anatomical descriptor of this standard view plane is the MAD measurement. It is, as BPD, the foundation of its standard view plane. At 18 weeks of gestational age, the MAD should be between 35 mm and 45 mm. A smaller or larger value implies that further action must be effectuated to investigate possible growth retardation. Because the MAD measurement consists of the average value of ATD and APD, both of these descriptors must be measured as well. It is essential that the two diameters are aligned perpendicularly and measured at the extremities of the abdomen.

The circularity of the abdomen is used to look for indications of a hernia containing abdominal contents. If the abdomen is compressed, it might be an indication of such a hernia. Both gastroschisis and omphalocele (see section 2.5.2) are associated with such findings.

The distance measurement related to the umbilical vein is included as an anatomical descriptor and the umbilical vein itself as an anatomical landmark. The reason is that this measurement is used to confirm a healthy fetus.

2.3.2.3 Medical basis

In this section, the medical reasons for choosing the landmarks of section 2.3.2.1 are given. In addition, it is discussed which abnormalities an irregularity in one or more of the landmarks might indicate. A listing of abnormalities and their associated symptoms present in the MAD standard view plane is presented in section 2.5.2.

Most landmarks of the MAD standard view plane are used merely to state that the view plane is obtained at the correct position. Its original definition was based on precise and easily obtainable landmarks, ensuring a common procedure for measuring the abdominal diameter. The umbilical vein was discovered to be the best reference point to use [10]. It is small, but easy to find at the same time, making it easier to accurately determine its position compared to a larger organ. In addition to the umbilical vein, a true transverse of the spine must be visible. To be sure that a true transverse is obtained, all three ossification centres of one vertebra should be visible and aligned in a triangular form. The third landmark, the symmetrical ribs must also be present to ensure that the correct section is made. If the ribs are not symmetrical, the view plane is obtained too much anterior or posterior.

The stomach bubble is included as a landmark for several reasons. Its presence confirms that the swallowing reflex is working correctly. If such a bubble is not present, it may be an indication of esophageal atresia. This abnormality occurs when the foregut does not separate into the ventral respiratory and the dorsal digestive tract [11]. The result will be a gap in the esophagus creating two separated tubes. The upper tube is connected to the mouth, while the lower tube is connected to the stomach.

If, on the other hand, two bubbles are present, one should suspect duodenal atresia (2.5.2.3). The two bubbles, often referred to as the double bubble, are the dilated stomach and a proximal duodenum. Duodenal atresia, including the double bubble is illustrated in figure 2.7.

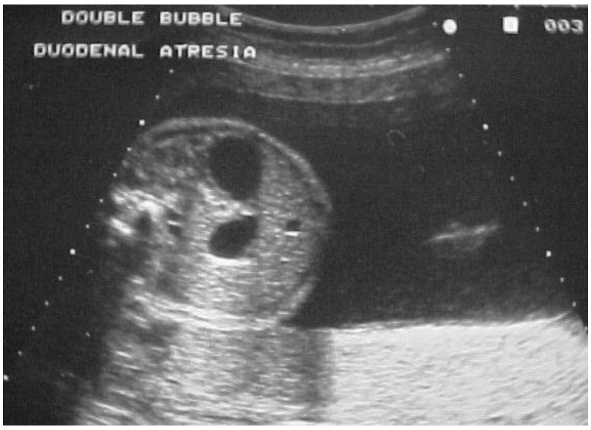


Figure 2.7: Illustration of the double bubble sign which is typical for duodenal atresia. Courtesy of [12].

During the obtainment of the abdominal diameter symptoms of other congenital abnormalities can be found. The ones presented in this thesis are gastroschisis (2.5.2.2), omphalocele (2.5.2.1), and diaphragmatic hernia (2.5.2.4). A MAD measurement deviating from the normal range should be the cause of further investigations. If the measurement is found to be decreased, diaphragmatic hernia should be suspected. This abnormality causes parts of the abdominal contents to protrude through the diaphragm into the chest cavity, causing a decreased abdominal diameter. If the abdomen has an irregular shape, causing difficulties when measuring the abdominal diameter, the fetus might have omphalocele, or gastroschisis. Both of these abnormalities cause the development of a hernia containing abdominal contents. The hernia is often attached to the abdominal wall, giving the outline of the abdomen an irregular appearance. These abnormalities and their symptoms are further elaborated in section 2.5.

2.4 Four chamber view (4CH)

Historically, the four chamber view was not obtained during the routine ultrasound examination at 18 weeks in Norway. At the end of the 1980s and beginning of the 1990s this changed making it an integrated part of the examination. The main reason for its incorporation is the frequent occurrence of cardiac abnormalities. According to Tegnander et al. [13], the incidence is about 8 in 1000 live births, making it the most common congenital anomaly. When it was discovered that a lot of these abnormalities could be found by examining the four chamber view of the heart, it became apparent that the four chamber view should be obtained at every routine examination.

The four chamber view does not provide any measurements, as is the case with BPD and MAD. It is used to visualise all four chambers of the heart in a single image. The image of the four chambers is taken horizontally through the fetal thorax, just above the diaphragm. It is essential that one entire rib can be seen to insure that the image is not oblique [14]. In addition to the four chambers, the septa and the valves are examined. The aortic arch and the ductal arch should also be visible in this standard view plane. An illustration of the four chamber standard view plane is shown in figure 2.8

Figure 2.8: Illustration of the four chamber view standard view plane.

2.4.1 Basis of choice

The four chamber view has been chosen as one of the standard view planes for several important reasons:

- Discovery of congenital heart defects
- Contains several distinct landmarks
- Integrated part of the usual examination
- Obtainable in nearly all fetuses

As mentioned above, the main medical reason for choosing the four chamber view as one of the standard view planes is the detectability of several congenital heart defects. Nyberg et al. [15] reports that 96 % of all sonographically detectable heart abnormalities can be found in this standard view plane. Examples of such abnormalities are septal defects and pulmonary obstruction [1] (for a complete listing see [15]). In addition to the abnormalities, the four chamber view is used to determine several other important aspects concerning the heart. Its size, location and orientation are determined, as well as size and continuity of the chambers and septa.

Another important reason for choosing this standard view plane is the fact that it contains several distinct landmarks which eases the task of accurately defining the it. The closed shape of the chambers makes it suitable for segmentation and further image processing.

Along with BPD and MAD, the four chamber view standard view plane is obtained at the routine examination at 18 weeks of gestational age in Norway. The same benefits apply here as they did for the BPD and the MAD standard view plane.

One of the main criteria for choosing a standard view plane is that it can be found regularly in nearly all fetuses. The study by Tegnander et al. [13] showed that the four chamber view could be obtained in about 96 % of all fetuses at the routine examination at 18 weeks. This makes it suitable for consideration. The drawback is that it requires some skill and training to obtain the correct view plane. The benefits of finding almost every congenital heart anomaly does however justify the additional training required.

2.4.2 Feature description

This section contains a description of the features needed to define the four chamber view standard view plane. The following sections describe the anatomical landmarks and anatomical descriptors.

2.4.2.1 Anatomical landmarks

The anatomical landmarks are used to define the standard view plane. In the four chamber view, the most important landmarks are:

- The atria
- The ventricles
- The valves
- The septa
- The aortic arch
- The ductal arch

The atria are the two upper cavities of the heart. These should appear equal in size and contain no irregularities in their walls. The ventricles are the two lower cavities of the heart.

When the four chamber view is obtained before 32 weeks, the ventricles should be of equal size (the right ventricle being larger at a later stage).

The valves control the blood flow within and out of the heart. There are a total of four valves. The atrioventricular valves prevent blood from flowing back into the atria during systole. The semilunar valves prevent the blood from flowing back into the ventricles during ventricular systole. In effect, the valves ensure the unidirectional flow of blood through the heart. The heart is a moving organ, and continuous opening and closing of the valves makes it hard to state a certain position in which they should be located. The user should be notified if the image is obtained when the valves are open.

The septa are thin structures separating the atria and the ventricles. These structures should be complete, without any ruptures. The septa form a cross with the atrioventricular valves, as the valves are positioned perpendicularly to the septa. It is important to notice that a small hole must be present in the atrial septum. This hole, called the foramen ovale, is present to ensure that the blood does not enter the lungs (which are non-functional prenatally). It can be seen as a small valve opening into the left atrium. A hole in the septum between the ventricles always indicates an abnormality.

The aortic arch, which supports the head and neck vessels can be seen as a candy cane shaped object entering the center of the heart. Further, the ductal arch is seen as a hockey stick shaped object directed towards the descending aorta [14].

2.4.2.2 Anatomical descriptors

This section will discuss the anatomical descriptors of the four chamber view standard view plane. The main descriptors are:

- A closed area and clearly defined regions
- Atria of equal size
- Ventricles of equal size
- An intact septa

In contrast to the two previous standard view planes described by this section, there are no quantitative measurements available when examining the four chamber view of the heart. The heart is a moving organ and any measurements being made, would be constantly changing. This forces the anatomical descriptors of this standard view plane to describe the static proportions between different parts of the heart.

A correctly obtained standard view plane of the four chambers of the heart should show four regions, one for each chamber. If the heart is healthy, these regions should be clearly defined, and possible to distinguish. The ventricles are clearly separated by the interventricular septum, while the foramen ovale makes the separation of the atria more difficult. Atrial septal defects can not be detected prenatally because of the foramen ovale.

After segmenting the four chambers, their relative size should be measured. There should be no significant differences in size when comparing the right and the left ventricle, and the right and the left atrium, and they should all have a smooth surface, containing no holes or irregularities. The main problem when creating these descriptors is the opening and closing of the valves. As there are no defined guidelines on when this standard view plane should be obtained, it is hard to ensure that the valves are closed in all images. A possible solution would be to notify the user when only two large chambers are obtained, i.e., open valves, causing the segmentation procedure to merge the ventricles and atria.

After describing the chambers, the septa must be examined. The interventricular septum is of particular importance, because it is intact prenatally, and should clearly separate the two ventricles. An indication of the presence of an interventricular septal defect can be found when trying to segment the ventricles. If the procedure fails, an incomplete septum could be suspected.

2.4.2.3 Medical basis

This section provides the medical background for choosing the anatomical landmarks stated in chapter 2.4.2.1. The landmarks are mainly used to be able to discover as many abnormalities as possible, in contrast to the two previous standard view planes, where the landmarks were used to determine correct obtainment of the standard view plane. This chapter defines which abnormalities each landmark might indicate by focusing on abnormalities where variations compared to a normal heart are large enough to be discovered during image analysis.

When analysing the four chambers of the heart, the relative size of the chambers should be carefully examined. There are mainly two abnormalities to look for, hypoplastic left- and right heart syndrome. When diagnosed with hypoplastic left heart syndrome, the left side of the fetal heart is underdeveloped. The reason for this abnormality is often a defect in the mitral valve. The blood does not flow into the left ventricle, causing hypoplasia. This means that if the size of the left ventricle is considerably smaller than the right one, hypoplastic left heart syndrome could be suspected. This should be confirmed by obtaining a long-axis view of the heart, to measure the length of the left ventricle and compare it to the total length of the heart [8]. Hypoplastic right heart syndrome is caused by defects in either of the valves at the right-hand side of the heart or an obstruction of the pulmonary artery. The only way for blood to enter the lungs is through the patent ductus arteriosus. The symptoms are the same as the ones of the hypoplastic left heart syndrome, only mirrored to the right side of the heart.

When examining the heart, the valves must, as mentioned in section 2.4.2.1, be examined. There are especially two abnormalities associated with pulmonary and tricuspid valve. If either of these are found to be stenotic, or even completely atretic the fetus is suffering from pulmonary or tricuspid valve stenosis[2]. This is a severe congenital abnormality in which the blood will not enter the lungs through the right ventricle as it should.

The interventricular septum should, as mentioned above, be examined to look for holes. A hole might indicate a ventricular septal defect, which is the most common congenital heart abnormality [8]. Although this abnormality is visible in the four chamber view, false dropouts may occur. This implies that further examination of the heart is needed if a ventricular septal defect is anticipated.

[2] Atresia is present if there is no blood flow across the valve

2.5 Abnormalities

The previous sections discussed the landmarks and the anatomical descriptors found in correctly obtained standard view planes of the fetus. Some abnormalities were also mentioned in the previous sections. These abnormalities will be further elaborated, and symptoms not located in the particular standard view plane are presented as well.

Abnormalities are usually not found by examining single images, but rather by moving the transducer through a larger area to look for several symptoms indicating a certain abnormality. This scope of this thesis is single image analysis, implying a more limited accuracy of the actual diagnosis. It will be assumed that additional information is present to increase the reliability of the diagnosis. Such additional information could include too much/little amniotic fluid, alcohol abuse by the mother, and diseases present at close relatives. Symptoms found in other standard view planes are also included to further strengthen the diagnosis.

When searching for a particular abnormality, certain symptoms are more indicative than others. While some symptoms confirm that the abnormality is present, others only state there is a possibility of the abnormality, but more information is needed to be certain. The indicative strengths of the symptoms are closely related to the actual presence of the symptoms in a certain abnormality. Most abnormalities are rather complex, which makes it hard to state exactly which symptoms should be present. A different approach is to state that some symptoms are always present, while others are occurring rarely, and possibly indicating other abnormalities as well. In order to categorise the symptoms according to the two factors just stated, each symptom is labeled as belonging to one of three categories: *may occur*, *quite frequent*, and *always present*. These categories are revisited when analysing the images of each standard view plane. Knowledge such as alcohol abuse or possible illnesses of the mother that may affect the fetus will not be categorised. They confirm the presence of an abnormality and will only be used when supporting a specific diagnosis.

2.5.1 BPD

This section will present some of the common abnormalities found in the BPD standard view plane. As mentioned by the previous section, the symptoms presented does not necessary belong to the BPD standard view plane, but could be located in other standard view planes of the fetal body. The abnormalities discussed are holoprosencephaly, hydranencephaly, anencephaly, spinal dysraphism, agenesis of the corpus callosum (ACC), and hydrocephalus.

2.5.1.1 Holoprosencephaly

Holoprosencephaly is, as mentioned in chapter 2.2.2.3, one of the most severe abnormalities affecting the cerebral hemispheres of the fetus. The most common symptoms of this abnormality are:

- Fusion of cerebral hemispheres (always)
- Fusion of the thalami (always (alobar), partial (semi lobar))
- Monoventricular cavity (always (alobar))

- Missing midline falx (always)
- Absent cavum septum pellucidum (always (alobar, semi lobar))
- Enlarged BPD (quite frequent)
- Polyhydramnios (may occur)
- Can be caused by: alcohol, maternal diabetes

As can be seen from this list of symptoms, all the symptoms (except polyhydramnios) can be found in the BPD standard view plane. This makes this abnormality rather easy to spot by using the information present in the BPD standard view plane.

2.5.1.2 Hydranencephaly

Hydranencephaly is another complex malformation affecting the fetal brain structures. The most common symptoms are:

- Absent brain structures (always)
- Partial or completely absent falx (quite frequent)
- Partial or completely absent septum pellucidum (quite frequent)
- Slightly enlarged BPD (may occur)
- Polyhydramnios (may occur)

As can be seen from the list, the first symptom is clearly the most indicative one. If an image appears with no brain structures present, hydranencephaly should be immediately suspected. Symptoms two, three, and four are all visible in the BPD standard view plane.

2.5.1.3 Anencephaly

Anencephaly is caused by a defect in the closure of the anterior neural tube [8]. This results in the following symptoms:

- Absence of skull (always)
- Myelomeningocele(s) in the lumbosacral or cervical region (quite frequent)
- Polyhydramnios (quite frequent)
- Can be caused by: folic acid deficiency, maternal diabetes

The distinct appearance of the absent skull makes it possible to diagnose this abnormality with a large degree of certainty using only knowledge obtainable from the BPD standard view plane. The second symptom of the list indicates the importance of having a thorough knowledge of the occurrence of related symptoms in different parts of the fetus when making a diagnosis. Without this knowledge, it is hard to make a reliable diagnosis.

2.5.1.4 Spinal dysraphism

Spinal dysraphism is a collective term for three abnormalities with a protrusion of neural elements and/or meninges through the spine. The abnormality is not visible in the BPD standard view plane, but several of its symptoms are, which justifies its presence. The most common symptoms of spinal dysraphism are:

- Protrusion of neural elements and possibly meninges through the spine (always)
- Lemon/bullet shaped head (quite frequent)
- Banana shaped cerebellum (quite frequent)
- Effaced cisterna magna (quite frequent)
- Dilation of bilateral, lateral and third ventricle (may occur)
- Can be caused by: maternal diabetes

As can be seen from this list, there are mainly two symptoms (second and fourth) present in the BPD standard view plane. This implies that whenever these signs are found, the examination of other standard view planes is required, to make the diagnosis of spinal dysraphism. The sagittal plane of the spine is a suggested place to start searching for protrusions.

2.5.1.5 Agenesis of the corpus callosum (ACC)

Agenesis of the corpus callosum is a serious congenital abnormality with symptoms present in the BPD standard view plane. The most common symptoms of this abnormality are:

- Absent corpus callosum (always)
- Largely separated lateral ventricles (always)
- Absent septum pellucidum (quite frequent)
- Local dilation of the occipital horns of the lateral ventricles (quite frequent)
- Cyst arising from the third ventricle (may occur)

Because the corpus callosum is not present in the BPD view plane, the second, third and fourth symptoms should be located before making any assumptions of this abnormality. ACC is especially assumed if the lateral ventricles are found to be separated.

2.5.1.6 Hydrocephalus

The last abnormality presented with symptoms in the BPD standard view plane is hydrocephalus. Hydrocephalus is a collective term used to describe an increased accumulation of cerebrospinal fluid in the ventricular system. It is caused by other congenital abnormalities with aqueductal stenosis and arnold-chiari malformation being the most common ones. This implies that the list of symptoms presented will include symptoms of these abnormalities as well.

- Dilated ventricles (always)
- Increased BPD (quite frequent)
- Protrusion of the cerebellum into the spinal canal (always (arnold-chiari malformation))
- Dilated proximal aqueduct of Sylvius (may occur (aqueductal stenosis))
- Asymmetric dilation of the lateral ventricles (may occur (aqueductal stenosis))

The symptoms visible in the BPD standard view plane are the dilated ventricles and the increased BPD measurement. Because this abnormality is caused by other serious congenital abnormalities, it is important to perform a thorough examination whenever any of these signs are found.

2.5.2 MAD

This section presents abnormalities found in the MAD standard view plane. The ones described are omphalocele, gastroschisis, duodenal atresia, and diaphragmatic hernia. Because of the late discovery (week 24 and later) of esophageal atresia, its symptoms will not be presented in any more detail than what was given in chapter 2.3.2.3.

2.5.2.1 Omphalocele

Omphalocele is a hernia connected to the umbilical ring containing abdominal viscera [8]. The hernia is surrounded by a transparent sac of amnion. It is one of the most common abnormalities related to the gastrointestinal system, with a rate of about one of 4000 births. The most common signs of this abnormality are:

- Cannot accurately measure MAD because the shape of the abdomen is malformed due to the hernia (always)
- Intestines in transparent sac of amnion attached to the umbilical ring (always)
- Polyhydramnios (quite frequent)

The most obvious sign of omphalocele visible in the MAD standard view plane is an irregularly shaped abdomen making the measurement of the MAD inaccurate. If such a finding is present either omphalocele or gastroschisis (see section 2.5.2.2 for a comparison) should be suspected.

2.5.2.2 Gastroschisis

Gastroschisis is an abnormality which have some similarities to omphalocele. While omphalocele is a hernia contained in a sac of amnion, gastroschisis is the formation of intestines outside the abdomen due to a paraumbilical wall defect. The most common signs of gastroschisis are:

- Cannot accurately measure MAD because the shape of the abdomen will be malformed due to the hernia (always)
- Bowel herniates through the anterior abdominal wall (always)
- Stomach malrotated (quite frequent)
- Can be caused by: low maternal age(<20), drug abuse

Because this abnormality is suspected whenever the MAD cannot be accurately measured, it is important to be able to distinguish it from omphalocele. It is, however, quite easy to separate them. Whenever the hernia is contained in a sac and attached to the umbilical ring it is omphalocele. If the hernia has an irregular shape indicating intestines without any protecting layer, it is gastroschisis. Gastroschisis is a herniation through the abdominal wall, and cannot be attached to the umbilical ring as is the case with omphalocele.

2.5.2.3 Duodenal atresia

Duodenal atresia is an abnormality with visible symptoms in the MAD standard view plane, characterised by a completely blocked duodenum. This causes a dilation of the stomach and the proximal duodenum will be visible creating the characteristic double bubble sign(see figure 2.7). The most common signs of this abnormality are:

- Double bubble (always)
- Polyhydramnios (quite frequent)
- Increased MAD (may occur)
- Can be caused by: maternal diabetes

This abnormality can, as seen from the list and by inspecting figure 2.7, be diagnosed directly by analysing the MAD standard view plane. The double bubble sign is in most cases enough to confirm the presence of duodenal atresia with a high degree of certainty.

2.5.2.4 Diaphragmatic hernia

The last abnormality described in relation to the MAD standard view plane is diaphragmatic hernia. Diaphragmatic hernia is a collective term for diaphragmatic defects where abdominal contents protrude into the chest cavity [8]. It is common to divide the abnormality into left- and right sided hernias, with the vast majority being left-sided. The list of symptoms will reflect this by only incorporating the symptoms visible in a left-sided hernia. An illustration of a left-sided hernia can be seen in figure 2.9.

- Decreased MAD (always)
- Heart shifted to the right side (always)
- Left part of the diaphragm is not visible (always)

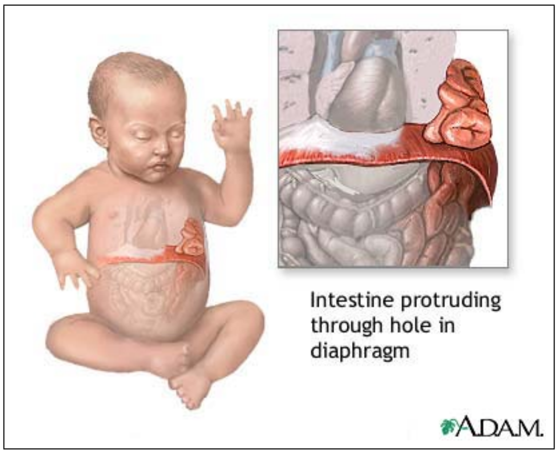


Figure 2.9: Illustration of a left-sided diaphragmatic hernia. Courtesy of [16].

- Stomach is in the chest (implying it will not be visible in the MAD standard view plane) (quite frequent)
- Displaced liver into the chest (may occur)

The only two symptoms found in the MAD standard view plane are the shifted heart and the displaced stomach. This implies a requirement of additional knowledge to discover this abnormality.

2.5.3 4CH

This section presents abnormalities found in the four chamber view standard view plane. The ones described are ventricular septal defects, pulmonary atresia, tricuspid valve atresia, and hypoplastic left heart syndrome. The reason for describing pulmonary/tricuspid valve atresia instead of stenosis is that a stenosis is much harder to discover, and often not found prenatally. Hypoplastic right heart syndrome was also mentioned in chapter 2.4.2.3, but because of its rareness compared to its left-sided equivalent, it is not included in this section.

2.5.3.1 Ventricular septal defects

Ventricular septal defects are by far the most common congenital heart defects, accounting for more than 20 % of all heart defects. The defect is, as mentioned in chapter 2.4.2.3, a hole in the septum separating the ventricles which causes oxygenated blood to flow from the left ventricle into the right ventricle, mixing it with deoxygenated blood. Being very inefficient, this forces the heart to pump more blood than it normally would. If the opening is large, this can cause the heart to be overworked and become enlarged. With only a small opening, the heart will function almost as normal, and such abnormalities are normally not found postnatally. They will even close by themselves in many cases. The most common signs of this defect are:

- Incomplete septum between the right and the left ventricle (always)
- Can be caused by maternal diabetes, alcohol

It is important to examine the heart in both the four chamber view plane and a view plane perpendicular to the septum to ensure an optimal visualisation of the septum. If only the four chamber view is used, false dropouts may occur. Beyond this, this abnormality is clearly visible in the four chamber view standard view plane.

2.5.3.2 Pulmonary atresia

Pulmonary atresia is a defect in which the pulmonary valve is absent or imperforate. The blood will stay in the right ventricle without flowing into the lungs. The only way for the blood to enter the lungs is by entering the right side of the heart through an atrial septal defect, and then enter the pulmonary vein through a patent ductus arteriosus. The most common signs to look for when investigating a possible case of pulmonary atresia are:

- Atrial septal defect (always)[3]
- Absent pulmonary valve (quite frequently)
- Small right heart (quite frequently)
- Obstruction of the pulmonary vein (may occur)
- Can be caused by: rubella

The obstruction of the pulmonary vein cannot be found by examining the four chamber view standard view plane. The remaining symptoms will, however, give a good indication of a possible pulmonary atresia requiring only a limited further investigation to make the correct diagnosis.

2.5.3.3 Tricuspid valve atresia

Tricuspid valve atresia is characterised by either a completely absent tricuspid valve (normally), or a tiny imperforate valve. The blood will not be able to flow into the right ventricle and hence not enter the lungs. This often causes defects in both the ventricular and the atrial septum allowing blood to flow from the right atrium into the left atrium, and further from the left ventricle into the right ventricle. The deoxygenated blood that was supposed to enter the lungs will then mix with the oxygenated blood, causing an insufficient supply of oxygen to the fetus. When trying to locate this defect, the most common signs to look for are:

- Atrial septal defect (always)
- Ventricular septal defect (always)
- Absent tricuspid valve (quite frequently)

[3]This defect is not possible to detect because of the foramen ovale, see section 2.4.2.1.

- Right ventricular hypoplasia (quite frequently)
- Pulmonary stenosis (quite frequently)
- Imperforate tiny tricuspid valve (may occur)
- Can be caused by: maternal diabetes mellitus

All symptoms except pulmonary stenosis should be present in the four chamber view standard view plane [4]. This indicates the possibility of confirming that the fetus has tricuspid valve atresia when obtaining the four chamber view standard view plane.

2.5.3.4 Hypoplastic left heart syndrome

The final abnormality discussed in this section is hypoplastic left heart syndrome. As mentioned in section 2.4.2.3, the left side of the heart will be largely underdeveloped due to an atretic mitral valve. The most common symptoms of hypoplastic left heart syndrome are:

- Left ventricular hypoplasia (always)
- Blood flows from pulmonary artery through ductus arteriosus to the aorta (always)
- Hypoplastic ascending aorta (quite frequently)
- Less developed valves (quite frequently)
- Reversed flow across atrial septum, from left to right (quite frequently)

The most obvious signs visible in the four chamber standard view plane are the hypoplastic left ventricle and the underdeveloped valves. To examine the blood flow from left to right atrium, a Doppler scan is needed.

2.6 Summary

In this chapter, the standard view planes of choice have been described. The ones chosen were BPD, MAD, and the 4CH standard view plane. For each of these standard view planes, anatomical landmarks and anatomical descriptors have been presented, and abnormalities commonly found have been listed. For further convenience, table 2.6 shows the abnormalities discussed in each of the standard view planes.

[4] A ventricular septal defect can be hard to recognise in this standard view plane, as mentioned in section 2.5.3.1.

Standard view plane	Abnormality
BPD	Holoprosencephaly
	Hydranencephaly
	Anencephaly
	Spinal dysraphism
	Agenesis of the corpus callosum
	Hydrocephalus
MAD	Omphalocele
	Gastroschisis
	Duodenal atresia
	Diaphragmatic hernia
4CH	Ventricular septal defect
	Tricuspid valve atresia
	Pulmonary atresia
	Hypoplastic left heart syndrome

Table 2.1: Abnormalities discussed in the standard view planes of choice.

Chapter 3
Feature extraction and understanding

This chapter will elaborate the second subtask presented in chapter 1, the extraction of features from ultrasound images. The first section of the chapter will describe and address the problems inherently added when using ultrasound and try to list their characteristics. The necessary properties of a possible solution will be derived from this listing in section 3.2. Section 3.3 describes some known solutions to the problem of segmentation. These solutions should fulfill most of the properties listed in section 3.2.

Section 3.4 describes the similarities and differences of automatic and manual segmentation routines. This section discusses the quality of a segmentation method provided by the two approaches, in order to create a foundation for choosing a particular method. Based on this discussion, the segmentation method of choice is selected.

After segmenting the regions of interest in the ultrasound image, a method to precisely describe them is required. The description is needed to match regions of interest with known organs. Section 3.5 describes several methods for describing a region, before making a conclusion as to which method should be used. The matching procedure is presented in section 3.6. Finally, a summary of the chapter is given in section 3.7.

3.1 Problem characterisation

This section will provide a characterisation of the problem at hand, which is to segment ultrasound images corrupted by noise. To understand the problem domain, a short introduction to ultrasound is given along with descriptions of the most common distortions.

Ultrasound waves are produced by applying alternating electric currents across the piezoelectric crystals of the transducer (the probe transmitting and receiving the ultrasound waves), which make them vibrate at a resonant frequency. This vibration is transmitted into the human tissue in small bursts. When the waves encounter the border between tissues of different densities (for instance soft tissue and bone), some of the waves are reflected back to the transducer. This makes the transducer act like a receiver, converting the mechanical waves to electric signals, which are used to display the ultrasound image on a video monitor. Because the velocity through human soft tissue is considered constant, it is possible to use the time it takes for the wave to return to calculate how far away from the transducer the reflecting object is located. This makes it possible to display the object at a proper location within the ultrasound image.

With the description of ultrasound provided by the previous paragraph, the main contributors producing noise in ultrasound images must be examined. Ultrasound imaging systems are based on a number of assumptions to be able to function properly:

- Constant speed of sound through homogenous tissue at 1540 m/s
- The propagation of ultrasound waves occurs in a straight line
- The ultrasound beam is infinitely thin in lateral and thickness direction
- The echo detected originates from the shortest path between the probe and the reflector and is produced by the last generated sound pulse
- The differences in echo amplitude and acoustic impedance between two adjacent tissues are proportional

These assumptions are, however, only theoretical constraints stated to make the development of ultrasound imaging systems possible. Practically, these constraints will not be applicable and as a result a number of artifacts may appear in the image. Angelsen [17] lists the following common artifacts:

- **Reverberation**, caused by inhomogeneities in the tissue which yield multiple reflections between tissue layers. These reflections will be added as a tail to the original pulse and produce ghost images at points with no real targets.
- **Geometric distortion**, caused by inhomogeneities in the tissue, yields spatial variations in the wave velocity. As the actual wave velocity will deviate from the expected velocity of 1540 m/s, problems related to depth ranging will occur. Objects that are assumed to be at a certain depth beneath the transducer are actually located either further away or closer to the transducer depending on the actual speed of the sound waves. Different wave velocities at the interface between two layers of tissue will in the same manner cause refractions which bend the beam and causes distortions in the image.
- **Wave front aberration** is an artifact caused by inhomogeneities in the tissue destroying wave fronts produced to steer and focus the beam. This aberration makes the beam diffuse and the spatial resolution of the resulting image is degraded.
- **The side lobe** is an artifact caused by the wave structure around the main lobe which will pick up signals from directions outside the image plane. These signals combined with interference appears as high frequency speckle in the image. The main lobe and side lobes are illustrated in figure 3.1.
- **Shadowing** is the loss of image data behind a dense structure as most of the sound energy has been reflected. This is caused by the reduced amplitude of the echoes reflected off a structure with high attenuation.
- **Enhancement** is the unexpected brightening of image data behind a translucent structure with low attenuation. This artifact is caused by echoes with increased amplitude due to the differences in attenuation of the adjacent tissue.

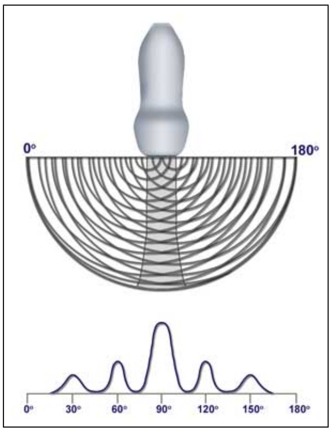


Figure 3.1: Main lobe and side lobes of ultrasound waves. Courtesy of [18].

3.2 Solution properties

This section will try to relate the constraints of ultrasound imaging presented in the previous section to the properties needed to successfully segment ultrasound images. These restrictions are inherently present when choosing ultrasound as the modality to use. To be able to utilise the information contained in such an image, the procedures used must manage the noise characteristics and try to remove or smooth the image without destroying or deleting valuable image data. The main challenge when developing a solution to this problem is how to distinguish between correct image data and noise or corruptions superimposed on the image.

The first important property of a solution for segmentation of ultrasound images is a robust initialisation. To ensure a high degree of reproducibility in noisy environments, the solution should be as independent as possible concerning the starting point of the segmentation. Being able to reproduce the same result from different initialisations not necessarily close to the sought object(s), increases the robustness of the system. A possible approach to avoid manual setting of initialisation parameters is the incorporation of meta data, such as transducer settings, in the segmentation procedure.

The next property required is the ability to distinguish between real edges in the image and false edges produced by noise or corruptions. False edges occur from distortions such as speckle noise and reverberations. The low resolution of the ultrasound image makes it quite sensitive to random speckle noise which may appear as part of an edge or even create a new edge. Noise and corruption can also cause real edges to be weakened and hard to detect during the segmentation. A solution for segmentation of ultrasound images should disregard false edges and detect edges that are weakened by noise.

When locating an object present in the image, it is important to be able to extract its contour for identification. Noise caused by for instance reverberations or signal attenuation might cause boundaries to be corrupted and discontinuous. A desired property of the chosen solution should be to correct this problem in such a manner that the true contour can be extracted.

The next section will describe some possible solutions satisfying one or more of the properties presented in this section.

3.3 Segmentation procedures

To be able to perform image analysis, it is essential to have a properly working segmentation routine. The analysis tool is dependent on segmented features to perform the analysis. If presented with an unprocessed image, it will not be possible to extract and combine the knowledge present in the image with domain knowledge.

There are several segmentation procedures currently used for segmentation of ultrasound images. This section will discuss some of these procedures, including their strengths and weaknesses. The discussion in this section forms the basis for the selection of the procedure presented by the next section. The evaluation is based on the results presented by the authors of the different procedures and recommendations given in [19]. The procedures presented are:

- Snakes
  - Attractable snakes
  - Gradient vector flow (GVF)
  - Early vision-based snake model
  - Robust active contour models (R-ACM)
- Combining snakes and active shape models
- Mathematical morphology
- Interacting multiple model probabilistic data association filter
- Live-wire

Each of these procedures are discussed in the subsequent sections.

3.3.1 Snakes

Snakes were introduced by Kaas et al. [20] in 1988. It has been used in several applications and has proven to be particularly useful when segmenting images in the medical domain.

The main idea of snakes is to try and fit an elastic band around the object of interest. This is accomplished by using internal forces which stem from the snake itself and external image forces. Using a combination of these forces ensures the possibility of exploiting both a priori knowledge about the structure at hand, by adjusting the internal forces, and the knowledge that is inherent in the image by adjusting the external forces. The internal image forces control the stretching and bending of the snake through two parameters, α and β. α is linked to the first derivative of the snake, enabling a discontinuous snake if set to zero. β is linked to the second order derivative of the snake, enabling the snake to develop corners if set to zero.

The external forces are derived from the image contents. There are several formulations of these forces, and some of the more common ones include:

- The actual image content
- Negative gradient
- Blurred image
- Blurred gradient

All of these forces are aligned in such a way that they attract the snake towards region boundaries. The accuracy of the external forces rely on its formulation, meaning that a blurred version of the gradient could improve the attraction of the snake. The increased complexity implied by having to adjust the parameters often limit the benefits of the forces.

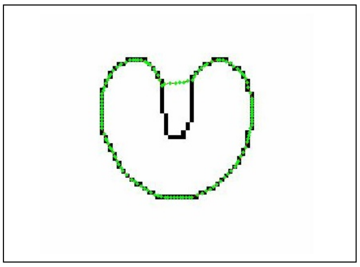


Figure 3.2: Inability to enter concave regions using snakes.

If the snakes were perfectly initialised with all the parameters correctly adjusted, it tempting to think that they are able to segment all objects. After all, they model a closed boundary incorporating both a priori knowledge and image knowledge. There are, however, some drawbacks to be handled. The main flaws of the original formulation were its inability to enter concave regions and the limited capture range, see figure 3.2 for an illustration of the problem concerning concave regions (for an illustration of the limited capture range, see figure 3.3(b)). These inabilities have led to alternative formulations of the snake algorithm. Two promising alternatives are attractable snakes [21] and gradient vector flow [22]. These approaches are explained in the two subsequent sections.

3.3.1.1 Attractable snakes

Attractable snakes were developed by Ji and Yan [21] in 2002. It is based on the greedy algorithm proposed by Williams and Shah [23]. The greedy algorithm has the same problems as the original snake formulation, and it is also prone to miss the contour of an object if the boundary is disconnected. The attractable snakes try to solve these limitations by using the concept of feedback control theory. A feedback energy term is added to the original energy formulation of the snake to control its movement. The idea is that the feedback term should control the movement of the snake in relation to both its position and its potential energy. Because of this dual control, the feedback energy term affects the snake by applying stronger forces to the snake when it is further away from the object, while decreasing the forces as the snake approaches the object. This increases the capture range of the snake, removing the initial limitation of the greedy algorithm.

In addition to the limitations of snake attraction, the original greedy algorithm also suffers from an unstable convergence criterion. The problem is that the snake might fail to reach an equilibrium, and therefore oscillate between two almost similar solutions [21]. This dramatically increases the time used to segment objects, causing the algorithm to be slower than necessary. The attractable snakes method uses a new and improved convergent criterion to prevent the solution from getting trapped in an oscillating state.

The movement of a snake is mainly determined based on three characteristics; the number of points that are moved from the last step ($M(V)$), the change of the snake's length ($\Delta L(V)$) and the image energy ($P_{image}(V)$). When the snake approaches equilibrium, $M(V)$, $\Delta L(V)$ and $\Delta P_{image}(V)$ approach zero, while $|P_{image}(V)| \gg 0$. By incorporating these three parameters into the convergent criterion, the snake will converge onto the desired contour even in the most complicated situations [21].

The main drawback of the attractable snake, is its sensitivity to noise. The original article by Ji and Yan only tests the algorithm on digital images and MR images, where the results are promising. As the work of this thesis is directed towards ultrasound images, the algorithm performance on similar images should be known. Abolmaesumia and Sirouspourb [24] apply the algorithm on ultrasound images, and report that the performance is not acceptable due to the high degree of speckle noise. The snake must be initialised close to the actual border, contrary to what is stated in the original article by Ji and Yan. This is caused by the increased amount of speckle noise which will prevent the snake from reaching the contour if initialised too far away.

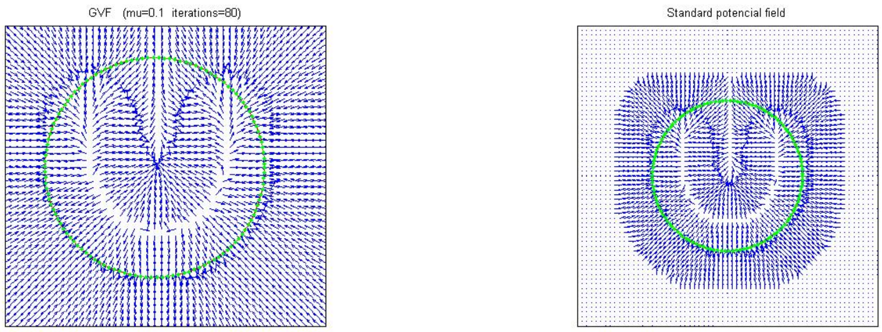
3.3.1.2 Gradient vector flow (GVF)

Another solution based on snakes is called gradient vector flow (GVF). GVF was first introduced by Xu and Prince [22] in 1998. The method aims at removing the same limitations as the attractable snakes. The solution adapted by Xu and Prince is quite different from the attractable snakes method. The main addition to the original method, is to extend the external forces. The strength of the forces are increased, implying and increased capture range of the object. To pull the snake into concave regions, a diffusion of the external force field is applied. This ensures that there are forces pulling the snake into the regions, instead bridging the gap without entering it.

An illustration of a normal force field and a GVF force field can be seen in figure 3.3. As can be seen from figure 3.3(a), the GVF field has a complete capture range. The snake can be initialised at the edge of the image, and will still be drawn towards the object. If the snake is initialised outside the force field of figure 3.3(b), it will not be pulled towards the object. Because of the internal forces of the snake, trying to keep it as regular as possible, it may eventually approach the object, but this requires more iterations than the snake using the GVF force field.

The arrows indicating the direction of the force field in the concave region in figure 3.3(a) are angled more vertically than the corresponding arrows in figure 3.3(b). This is due to the diffusion of the force field, ensuring that the snake is pulled into the cavity.

The main limitations of the GVF implementation are common for all snake implementations. The snake is easily trapped at spurious edges or artifacts. In the idealised world of figure



(a) Illustration of the GVF field (b) Illustration of the standard potential field

Figure 3.3: Comparison of the GVF field and the standard potential field.

3.3 it is not a problem, but when working with ultrasound images, this is a major drawback. Another disadvantage is that objects tend to be grouped close together, enforcing a very accurate initialisation, limiting the advantage of the increased capture range. Snakes are also heavily dependent on a correct parameter adjustment. In addition to the parameters α and β of the original snake formulation, GVF incorporates two more weighting parameters to control the external forces. These values must often be tuned to each specific case, causing difficulties when trying to create a fully automatic segmentation procedure.

3.3.1.3 Early vision-based snake model

The early vision-based snake model was proposed in 2000 by Chen et al. [25]. The method was developed to solve the limitations present in the general snake formulation, but also limitations of the GVF model mentioned in the previous section. The implementation differs from the original snake formulation in two ways. First, the energy minimisation used to move the snake is created using the distance map of the original image. The other difference is that each snaxel[1] is moved discretely. The advantages are reduced influence by noise, and a better image force because of the highlighting of the edges [25]. In addition, the discretisation greatly reduces the search space for the movement of each snaxel. The search will be done within the local maxima instead of searching through every pixel along the search path.

The problem of adjusting the weights of the internal and external forces has also been handled in the approach by Chen et al. The weights are changed to maintain a certain ratio between the four forces (two internal and two external). This ensures that the weights are changed in relation to the underlying image boundaries.

A comparison between this method and the GVF method is given in figure 3.4 (for a more complete comparison, see the original article by Chen et al. [25]). The improved performance of this method compared to GVF is evident when examining this figure. Even when the snake is initialised closer to the object at hand using the GVF method, it is largely outperformed by the method by Chen et al.

[1] Snaxel is a discrete point in a 2D snake.

(a) Segmentation by GVF (b) Manual segmentation (c) Segmentation by method proposed by Chen et al.

Figure 3.4: Comparison of segmentation results of the GVF method, a manual segmentation, and the approach by Chen et al. Adapted from [25].

3.3.1.4 Robust active contour models (R-ACM)

The robust active contour model (R-ACM) was proposed by Svinning [19] in 2003. It is based on the approach presented by Chen et al. [26]. To fully understand R-ACM, the approach by Chen et al. will be introduced. This method uses the discrete snake presented in the previous section, but instead of using a single snake, two snakes are used, one inner and one outer. The inner snake being an internal point of the object at hand, and the outer snake surrounding the object (this method implies knowing the size of the largest object to segment). The movement of each snaxel is defined along a path between these two contours, causing each snaxel to converge at a final point when reaching equilibrium. To avoid the snake getting trapped at local minima (typically speckle noise in ultrasound images), which is a problem using the original discrete formulation, the notion of discrete gradient flow (DGF) is introduced. The gradient flow creates a smooth transition between two local maxima by first performing edge detection on the path between the snaxels on the inner and outer snake. The magnitude of each edge is then computed, and only the local maxima are kept. The final DGF will therefore contain all the local maxima and the global maximum along the path. The magnitude of the edge between two local maxima is simply a linear interpolation of the magnitude of the maxima. By applying the interpolation scheme between the maxima, a continuous path between all inner and outer snaxels where the magnitude has been computed is ensured. The inverse (to ensure a minimization of the energy) of the magnitude is used as the external force which is applied to both the inner and outer snake.

In addition to the DGF force, a new force is added to further prevent the snake from halting at insignificant edges. The force, called a DS-potential force, is used to pull the two snakes together. It is introduced by letting the most stable snake pull the other snake towards itself, causing it to cross spurious edges and speckle noise. It is, however, difficult to adjust the weights of the DS-potential force, i.e., how hard the two snakes should pull on each other. To solve this, a three-step procedure is proposed. First, only the inner forces with the DGF force is used, creating an initial position for both snakes. When the snakes have settled down, the DS-potential force is applied to further adjust the position of the snakes, pulling them together to form a single snake. The third step is optional, and includes a smoothing of the final snake.

Because of the limitations still present, Svinning extends this method by changing the weighting scheme of the forces and introducing a new stage between the first and the second step described above. The problems of the method proposed by Chen et al. [26] arise when one of two conditions occur. First of all, the method does not perform well when there is missing edge information. When adjusting the weights in this scenario, the snake believes it is dealing with two separate edges causing it to move towards one of these edges while it instead simply should have bridged the gap. The second problem occurs when the object to be segmented contains a more complicated topology than a circular or rectangular shape. If the path between two snaxels creating the DGF crosses two edges, the DGF will contain two maxima, and it can not be determined which is the correct edge. The advantage of deploying the initial outer contour far away is hence removed.

The improved weighting scheme makes sure that objects far away from the contour or objects with size different from the object at hand is given less significance. They will therefore not be able to attract the snake as much as the closer, and proper-sized objects. The problem of missing edge information is also handled by relaxing the criterion of the snaxel movement. It enables the snake to be controlled by the internal energy when a gap is present. Since the internal energy aims at having a smooth surface of the snake, the gap will be closed, preventing the snake from being attracted to other objects.

The main extension made by Svinning is however the inclusion of the new stage between stages one and two in the previously described method. Whenever the DGF path crosses the boundary of the object more than once, the original approach fails. Svinning has a method to solve this by splitting the regions between the inner and outer contour into more regions, and reparameterise the contours. New DGF paths are created within these regions which guide the snake along the correct boundaries of the object. This process must be completed until no further splitting of the regions can be performed (i.e., no new regions can be found.

The results presented in [19] indicate some good preliminary results using R-ACM. We feel however that the results do not improve significantly compared to the results presented in for instance [26]. The reason for this might be the choices made concerning the ultrasound images which are segmented.

3.3.2 Combining snakes and active shape models

Section 3.3.1 explained the application of snakes in general, and several improvements of this method in detail. An interesting approach proposed by Hamarneh and Gustavsson [27] combines snakes and active shape models[2] when segmenting the left ventricle in ultrasound images.

The idea of this method is to have a training set of manually segmented images. This training set is used to represent the normal variation of the object to be segmented. The solution used in [27] is to first apply a discrete cosine transform to all the training images, and then perform what is called a principal component analysis (PCA) [28] to identify the main modes of variation. By performing PCA in the frequency space, you eliminate the need for point correspondence. After performing PCA, the snake is initialised and allowed to deform once using traditional snake forces, as described in section 3.3.1. To see whether or not the snake

[2] An active shape model tries to model the natural variations of the object to be segmented.

has moved within its allowed space, as defined by the main modes of variation, the snake is transformed into the frequency domain and projected into the allowable space. The snake is then transformed back into the spatial domain, and this process is repeated until the snake stabilises, i.e., most of the nodes of the snake have reached their equilibrium.

The main advantage of this method is the incorporation of the a priori knowledge about the shape of the object. Knowing this in advance, makes it possible to delimit the movement of the snake, and hence avoid it being dragged towards distant objects. This feature may also be seen as the major drawback of this method. It requires a lot of work to manually create the training set for each object, and it is not suited as a generic segmentation procedure, because of its specialised nature.

3.3.3 Mathematical morphology

Mathematical morphology is a general image processing procedure which can be used at several stages of the image processing pipeline. In contrast to other procedures, morphology is based on the shape of objects instead of the relative intensities of the pixels. This makes it especially suited for the processing of images where features of a certain size should be removed or added, using the elementary operators dilation and erosion. Other high-level procedures can be developed by combining these two operators.

A promising example using mathematical morphology was proposed by Thomas et al. [29]. By using a pipeline of different morphological procedures, the femur bone is segmented and measured. The results presented are quite promising, but there are a few weaknesses that should be commented on. First of all, because of the large amount of steps in the pipeline, the setting of parameters (size of the structure elements) must be tuned to the specific object. This makes it unsuitable for generic usage. Secondly, morphological methods are quite slow, especially when working with large images. An increase in the size of the structure element will further increase the amount of time for the procedure to complete. These limitations, in addition to the fact that very few new approaches have been developed during the last years, have led to the conclusion that morphology is not a solution when a robust and generic segmentation procedure is required.

3.3.4 Interacting multiple model probabilistic data association filters

Another interesting technique for segmenting ultrasound images was proposed by Abolmaesumi and Sirouspour [24] in 2004. The authors have identified the need for a segmentation method working without human intervention (initialisation) and also without the slow convergence criterion present in for instance snake algorithms. Their idea is to create the boundary of the object based on the path of an object travelling along a path at given radii from a seed point inside the object. This path is governed by a finite set of models at any radius. By applying a Kalman filter [30] to each of the models, they can be used in the recursive IMM/PDAF [24] algorithm to estimate the location of the prostate boundary.

Three steps are performed to extract the actual boundary. First, the initial state estimates used by the Kalman filters are calculated. These estimates include the transition probabilities, i.e., the chance of making a switch from model i at step $k-1$ to model j at step k. The next step performed is the mode-conditioned filtering step. The PDA filter is used to calculate the mode-conditioned state estimates and the covariances of the next iteration. These figures are used in the final step of the method when predicting the boundary of the object.

The results presented by Abolmaesumi and Sirouspour show a satisfying segmentation of objects in ultrasound images. A comparison to manual segmentation shows only small differences in the area of the objects and maximal shortest distance, implying the obtainment of a good match. The algorithm is fast compared to the traditional segmentation procedures presented earlier in this section.

Since the boundary points are all placed on the path at a given radius from the seed point, this point must be placed so it can reach all the boundary points without crossing any region borders. This is probably the main limitation of the method, as it requires that the object has a certain topology. Another drawback of the seed point is that it offers low reproducibility. The change in positioning of the seed point will also change the resulting boundary.

3.3.5 Live-wire

As a fully automatic segmentation procedure of ultrasound images is often very complex or even impossible (if a high quality procedure is required), a semi-automatic approach could be more affordable. A possible semi-automatic solution is the live-wire approach, a contour-based segmentation technique presented by Mortensen et al. [31]. The algorithm is based on graph searching where the image pixels are treated as nodes and the costs between neighboring pixels as edges in the graph. Dijkstra's path searching algorithm is then used to extract the final contour. Using live-wire, a contour of a target object is constructed by interactively selecting control points which should be members of the final contour. The actual extraction of the contour is based on calculating the minimum cost paths between a seed point and the currently selected point. The costs associated with each point is calculated as the image's gradient magnitude and gradient orientation. The final contour yields an optimal path in terms of minimum cumulative cost.

The live-wire technique yields a semi-automatic approach that decreases the boundary extraction time and accuracy compared to a pure manual segmentation, while reproducibility is increased. Ji et al. [32] shows some examples of using live-wire segmentation applied to MR-images. Compared to a fully-automatic approach, one of the drawbacks of the live-wire is the actual user interaction needed to be able to establish a seed point. Another drawback is the potentially large graph searching area, which depends on the size of the image being segmented.

3.4 Automatic or manual segmentation

This section will elaborate the differences and similarities between automatic and manual segmentation of ultrasound images. The intention is to confirm whether or not the methods developed for automatic segmentation are able to produce results of the same quality as the manual approach. As the actual segmentation is not the main area of interest in this thesis, it is important to establish how much effort should be invested in the subject.

The automatic segmentation of ultrasound images is still an area of research. No general approach applicable for ultrasound images is yet known. Progress is found in several areas of interest, with the production of synthetic results [33], [34].

Most automatic segmentation methods are often semi-automatic in some aspects. They often require an initialisation to perform within acceptable limits. On the other hand, a semi-automatic approach will enable the user to verify each frame to gain greater confidence in the final result. A manual initialisation will narrow the solution search space significantly, but requires expert knowledge on the object of interest. The main characteristics often linked to automatic or semi-automatic segmentation are:

- High reproducibility
- Variable noise sensitivity
- Highly dependent on image data
- Often not fully automatic

Manual segmentation is the process of drawing all the contours required by hand, using some kind of pointing device. This is highly time consuming if there are series of images or complex structures, requiring expert knowledge. Using manual segmentation as part of a system for image analysis will create a substantial processing bottleneck and a low degree of reproducibility. The characteristics of manual segmentation are:

- Variable reproducibility
- Incorporates multiple area of knowledge
- Requires expert knowledge
- Requires less pre-processing
- Time consuming

When processing and analysing ultrasound images, the amount of time needed for a complete manual segmentation is not affordable. The main focus of this thesis is the analysis of the contents of the standard view planes, making it reasonable to use a semi-automatic segmentation routine. This ensures a robust segmentation (due to its automatic nature), while at the same time incorporating domain knowledge (due to the expert guiding the procedure). These requirements are fulfilled by the live-wire approach presented in section 3.3.5. More details on live-wire can be found in appendix A.

3.5 Region description

The previous sections discussed the process of segmenting the regions of interest in the standard view planes. This is, however, not enough to proceed to the next step in the pipeline, which is the analysis of the contents of the ultrasound images. The segmentation process offers a way of obtaining the region, without stating what it should represent. This means that a method of identifying the organs present must be developed. The first step of this process is to make a thorough description of the segmented regions. If known organs are already described in a similar way, a matching procedure can be developed to label each region according to the most similar organ.

The subsequent sections describe several approaches to the task of region description. When discussing the different approaches, it is important to keep in mind which domain the procedure is situated in. The fact that this thesis concerns medical images, and especially ultrasound images imposes some requirements as to which methods are more appropriate than others.

3.5.1 Chain code

A common way of representing regions, by using chain codes. They describe a region by tracing the border of a region and encoding it according to the orientation of the step between adjacent pixels. An example is shown in figure 3.5. Because of its simple initialisation, chain

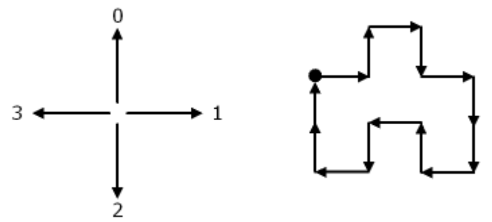


Figure 3.5: Illustration of a chain code. This region will have a chain code of 1,0,1,2,1,2,2,3, 0,3,2,3,0,0.

codes can be generated whenever the object is segmented, which is its main benefit It is, however, not very suitable for representing arbitrary objects. Since every small deviation along the contour is encoded, the chain code is extremely sensitive to noise. This renders this method quite useless in the domain of ultrasound images.

3.5.2 Fourier descriptors

Fourier descriptors describe the shape of an object in the complex plane. When the region is transformed using the Fourier transform, the Fourier descriptors, T_n are given according to the following equation:

$$T_n = \frac{1}{L}\int_0^L z(s)e^{-i(2\pi/L)ns}ds \tag{3.1}$$

where L is the curve length of the object.

The Fourier descriptors contain all the information about the object, and this is obtained by transforming the object back into the spatial domain. The reason Fourier descriptors are used to describe regions, is because only a few of the descriptors are needed to maintain the outline of the object. The lower order descriptors contain low frequency information, while the higher order descriptors contain high frequency information. This means that if it is not of particular interest to describe every detail of the region, the high frequency components can be discarded, and hence reduce the amount of information needed. This is illustrated in figure 3.6, where the original region is fully defined by 64 descriptors, but the outline is still kept when using only 8 lower order descriptors. The shape does not change radically before 62 descriptors are used. When working with large datasets, the reduction of the information could be of high importance.

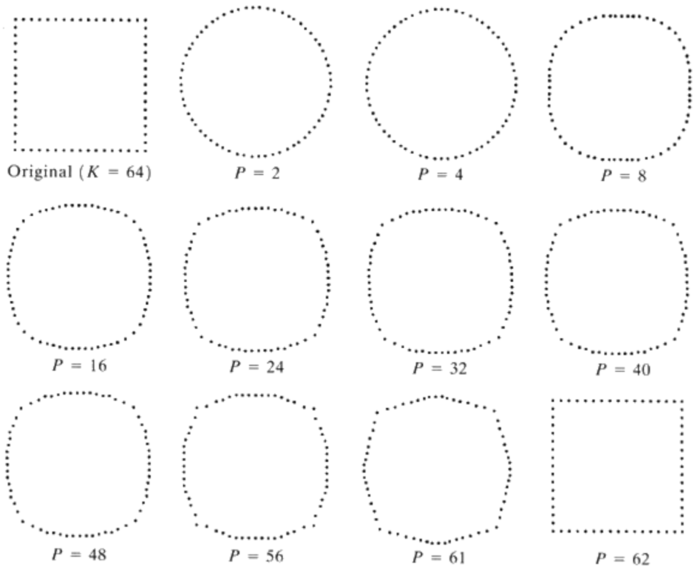


Figure 3.6: Illustration of reconstruction of the shape using P descriptors. Courtesy of [35].

The main problem concerning this method, is the fact that a lot of the high frequency components must be kept to correctly define each region. There are often small details that separate a healthy organ from an organ containing an abnormality. If the difference is removed during the description phase, the final diagnosis would be unreliable.

3.5.3 Region-based descriptors

The descriptors presented by the previous sections are classified as contour-based descriptors. This section presents a set of descriptors based on the whole region, region-based descriptors.

Before discussing the descriptors chosen by this thesis, the concept of invariance will be introduced. A descriptor is said to be invariant under a geometric transformation, if the transformation of the object does not change the numeric value of the descriptor.

When working with ultrasound images, the amount of invariance needed is defined according to the variations present in the image. First of all, the resolution of the images often change according to the level of detail needed. This means that the descriptors should be scale-invariant to incorporate the different resolution present in similar standard view planes. In addition, because the fetus does not have a predefined size, the different organs will vary in size, increasing the importance of having scale-invariant descriptors. Secondly, the fetus is constantly moving, and it is hard to tell where a certain organ should be. Since the movement is not restricted to any dimension, the organs can be both translated and rotated according to the original definition. This implies the need for both translation- and rotation-invariance.

This thesis focuses on essentially four region-based descriptors, which are area, circumference, compactness, and statistical moments. This section discusses their degree of invariance in order to state their appropriateness in the medical domain. Area and circumference are naturally invariant to rotation and translation, since they are simply a numeric value representing the size of the object. If the area (and circumference) is defined to be the ratio between the number of pixels constituting the region (border) and the number of pixels in the image, they will both be scale-invariant. Since compactness is a combination of area and circumference, it is also invariant to both rotation and translation (scaling if the area and circumference is defined to be scale-invariant). The statistical moments, however, can be made invariant under all three transformations. Sonka et al. [36] present a list of seven statistical moments which are all invariant under the three transformations needed in the medical domain. The mathematical foundation for the region-based descriptors are found in appendix B.

The main advantage of region-based descriptors is that they have a large degree of invariance to geometrical transformation of the region. This is, as mentioned above, highly important when trying to uniquely describe organs. The descriptors are also independent of the chosen segmentation routine. They describe the general shape of the object, and this is not changed dramatically if a different segmentation routine is being used[3]. The fact that these descriptors create a numerical vector describing the object makes them highly suitable for later matching against other objects. A simple comparison can then be made in order to determine the most similar object. A summary of the region descriptors presented and their degree of invariance is found in table 3.5.3.

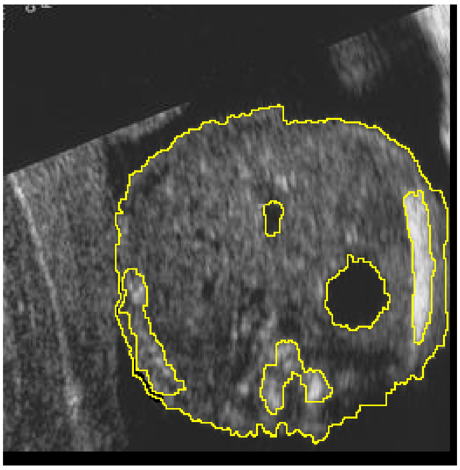
Descriptor	**Invariant to**
Area	Rotation and translation (scale)
Circumference	Rotation and translation (scale)
Compactness	Rotation and translation (scale)
Statistical moments	Rotation, scale, and translation

Table 3.1: Descriptors and their degree of invariance.

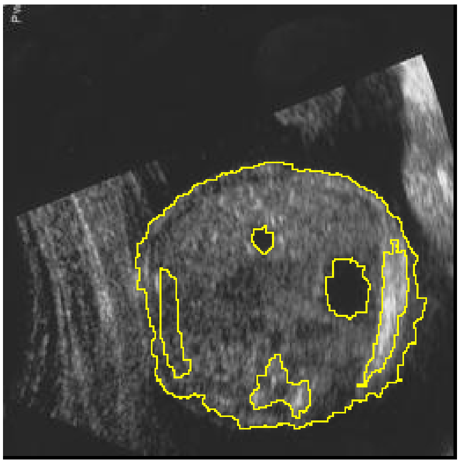
3.6 Region matching

The previous section presented a method for describing each region in the standard view plane. This is, however, not enough to perform a proper analysis. The description of each region must be used to identify the organ it represents. This implies that a description of known organs in the standard view planes must be provided. By describing known organs in a similar way as presented above, a matching procedure can be performed between known

[3] In contrast to for instance the chain code, which is highly dependent of the segmentation routine.



(a) Labeled image



(b) Unlabeled image

Figure 3.7: Images used as input for the matching.

organs and regions obtained and described from the standard view plane. This means that the difficult task of identifying the organs in the image, is reduced to finding the most similar description in an already described set of known organs.

The matching procedure is performed by calculating the total deviation between the descriptors of each region. This gives a score stating the degree of match between the unknown region and all the described organs[4]. The lowest score indicates the most probable match, and the unknown region is given the same label as the known organ.

The results of a matching procedure between the objects of two images are presented in table 3.6. The left column of the table indicates the numbers given to the regions of the unlabeled image[5], while the first row indicates already described organs of the same standard view plane. To illustrate this, both the unlabeled image and labeled image are included in figure 3.7. When comparing the table and the images, the best score do indicate the same organ in the two images.

	mad_measurement	rib_right	umb_vein	stomach_bubble	rib_left	spine
1	**0.066**	1.5876	6.1296	1.2270	1.4598	1.3317
2	0.2303	0.5943	**0.1323**	0.1900	0.5524	0.3962
3	1.1343	**0.1099**	1.3949	1.1918	0.1864	0.6055
4	0.2227	0.4867	0.3603	**0.0822**	0.5248	0.3082
5	0.9891	0.2951	1.0324	1.0877	**0.0818**	0.5526
6	0.2734	0.4799	0.4198	0.2495	0.4175	**0.1489**

Table 3.2: Computed score between the regions of the images in figure 3.7.

If the lowest score of a particular region is above a predefined threshold, the region is marked

[4]The matching procedure is only performed against organs known to exist in the particular standard view plane.

[5]The number of each region is determined top to bottom, i.e., the topmost region is number one, etc.

as unknown. Whenever unknown organs are found, two scenarios may have happened. Either, the region found is previously not described, or the region is too dissimilar to the description of the organ it actually represents. Both of these scenarios require the consultation of a medical expert before proceeding.

After recognising which organs are present in the standard view plane, the descriptors presented in chapter 2 are calculated. They are added to the total knowledge describing the contents of the particular standard view plane. Without combining this knowledge, it is not possible to analyse the contents of the standard view plane.

3.7 Summary

When trying to segment ultrasound images one is faced with several new challenges not present in other images. Because of the high degree of noise, the segmentation procedure has to be very robust to obtain good performance.

This chapter presented the challenges inherent in ultrasound images, before solution properties were described to solve these challenges. Section 3.3 presented several methods incorporating one or more of these properties that are currently being used for segmentation of ultrasound images. The one found to show the most promising results was the R-ACM method proposed by Svinning [19]. Although, as stated earlier, the results presented did not show an evident improvement compared to the dual-snake model proposed by Chen et al. [26], the improvements in the weighting scheme and its ability to handle complex topology makes it a suitable choice if automatic segmentation of ultrasound images should be implemented. Because the main focus of this report is the analysis of images, and not segmentation, existing automatic segmentation methods are only described. As mentioned in section 3.4, the live-wire approach presented in section 3.3.5 will be used when segmenting the images of each standard view plane. A more complete mathematical definition of this method is given in appendix A.

Based on the discussion of the region descriptors in section 3.5, the region-based approach has been found most suitable for describing the objects in the standard view planes. They fulfill the requirements concerning the movement of the fetus, while at the same time offering an easy way of matching against similar objects. Both chain codes and Fourier descriptors have limitations making them quite unsuitable for describing the regions extracted from ultrasound images.

After obtaining the description of each region in the standard view plane, the regions must be recognised as being specific organs. This is done by matching the region descriptors obtained with descriptions of known organs. The result of this process is that a label is assigned to every region, stating the organ it is most likely to represent. If a good match is not found[6], the region is labeled as unknown, requiring consultation of a medical expert before proceeding.

The final requirement before commencing the analysis of the images, is the extraction of the descriptors presented in chapter 2. These descriptors are combined with the organ labels to complete the knowledge case describing the contents of the particular standard view plane.

[6] Lowest match is above a predefined threshold

CHAPTER 4

FEATURE ANALYSIS

This chapter will elaborate the third subtask presented in chapter 1, the analysis of the extracted features. The analysis will be based on the knowledge gained through the feature extraction and understanding presented in chapter 3 and medical domain knowledge. The concept of feature analysis is elaborated in section 4.1. Section 4.2 presents a definition of the concepts data, information, and knowledge. It is important to have an understanding of these three concepts when discussing feature analysis. The rest of the chapter will explain and discuss two methods for performing the analysis. Case-based reasoning is elaborated in section 4.3, while neural networks are presented in section 4.4. After presenting the methods, their differences are discussed in section 4.5, before a conclusion is drawn in section 4.6.

4.1 What is feature analysis

Conceptually, feature analysis is the process of combining the knowledge inherent in the image with additional domain knowledge to reach a conclusion about the contents of the image. The range of the analysis can be from simple procedures like scanning bar codes getting the price of a product, to more sophisticated methods such as automatic face recognition. The main difference between these two methods is the variance in the data presented. When scanning a bar code, the amount of variance present in the image is very small compared to the original image. The main objective of identifying clearly separated lines and measure their relative spacing and size is a well-formulated task requiring little processing. If the task of the analysis is to recognise a face, the amount of variance in the input image is much higher because the method should recognise different facial expressions of the same face. This is an easy task for people to perform, but when only presented with a large matrix of pixel values, the task is much more complicated. In order to solve such a task, it is essential to discover landmarks, or essential features, in the image. Using these landmarks it its possible to separate one face from another and also to correctly identify the same face with different expressions.

It should be clear that the task of analysing ultrasound images to look for possible abnormalities, or, hopefully to determine that the fetus is healthy, belongs to the latter of the two tasks presented above. The landmarks will be extracted using the segmentation procedures as presented in section 3.3. It is, however, not enough just segmenting these features. To be able to do an analysis of the image contents it is essential to know the meaning of, and the mutual relations between the objects. This is especially important when working with medical image contents, since a symptom at one location will often suggest that something is wrong in other parts of the body. In order to incorporate this extra knowledge into the image analysis process, a domain knowledge model is required, describing the relations between objects, and the possible implications inferred by specific findings. This makes it possible to determine the consequences of the findings, and also to relate the findings to each other. By creating the domain knowledge model, it is also possible to take advantage of additional information not present in the image when performing the image analysis. It is known that there are several factors affecting the chances of a fetus being healthy at birth. Examples of such factors are a smoking mother, and diseases which are found in close relatives. By using this information it is possible to get a more accurate picture of the situation of the fetus, which in turn will lead to a more reliable analysis and diagnosis. The domain knowledge model and its contents will be further explained in chapter 5.

When discussing the two primary methods for performing the image analysis task, it will be considered how easy it is to incorporate the domain knowledge model and its knowledge into the task. This will be one of the main criteria when choosing the method in section 4.6.

4.2 Data - Information - Knowledge

To be able to clearly separate the concepts of data, information and knowledge, this section will try to explain their properties and relations. Figure 4.1 illustrates the connection between these terms.

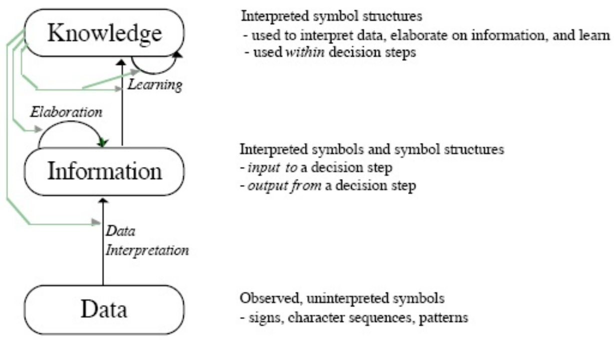


Figure 4.1: The relations between information, data and knowledge. Adapted from [37].

Data are simple and syntactic entities which on their own provide no meaning to the subject of interest. Examples are single characters, symbols, signs, patterns etc. To extract information from data, it must be given meaning by some interpretation. Depending on the subject of reference, different meaning can be interpreted from the same data. The process of interpretation extracts meaningful information from the data using knowledge of the particular domain.

The output from the previously described data interpretation is information. This information might be the initial interpretation of the data and might require further elaboration in order to get a better understanding. As shown in figure 4.1, the information is both the input and output of an elaboration process. After being sufficiently elaborated the information is given as input to a learning process.

The information submitted to the learning process is transformed to knowledge when it has been processed and incorporated as part of the reasoning resources of the subject of interest. Learning is often considered to be the process of integrating new knowledge into already existing structures in such a way that it can be utilised for making decisions at a later stage. As shown in figure 4.1, the knowledge learned is later used to interpret data to information, elaborate and possibly derive new meaning from information, and learn from new knowledge.

4.3 Case-based reasoning(CBR)

This section will elaborate the principles and properties of CBR along with its usage in image analysis and more specifically the medical domain. Section 4.3.1 will explain the basic properties of CBR, its fundamentals and its cycle of phases. This section is influenced by the work of Aamodt([38], [37], [39], [40] and [41]). Section 4.3.2 will give a description of three CBR systems using different approaches. Strengths and weaknesses of the CBR paradigm will be described in section 4.3.3. Complex environments and domains such as the medical domain gives raise to a lot of problems as far as knowledge is concerned. Section 4.3.4 will focus on the most important ones.

4.3.1 Properties of CBR

Case-based reasoning is a paradigm of problem solving where specific knowledge about previously experienced problem situations are utilised together with general knowledge of a problem domain. A new problem is solved by finding a previous similar case and match its description with that of the current. The solution of the past case is then reused in the new case. The reasoner learns from each new problem solving session by retaining the relevant information from the currently solved case and makes it available for further problem solving.

The most important steps of CBR are finding a good match to a new problem, adapt a previous solution to successfully solve the new problem and decide how to index and store cases for effective retrieval later.

The underlying assumption of CBR is that similar problems often have similar solutions. This is the same way the human brain examines a new problem with an unknown solution. If there is a way to apply a previously used solution, it is the most effective approach for solving the new problem.

CBR methods are often divided along the data-knowledge dimension. Data-intensive and knowledge-poor methods consider cases as data records and assesses similarity based on a simple metric during comparison. A knowledge-intensive, but data-poor method considers a case as a more complex user experience and uses an explanation process during the similarity assessment. A combined CBR method being both knowledge- and data-intensive will utilise multiple similarity assessment methods as well as multiple case contents. A knowledge-poor method is used when general domain knowledge is hard or impossible to acquire.

In the medical domain, knowledge is generally present in two ways: as rules in textbooks and inherent as experience of experts. The knowledge of the expert is often organised in cases, and matters concerning specific patients and diseases that are previously handled. Schmidt et al. [42] divides the knowledge of the medical domain into objective knowledge, that of the textbooks, and subjective knowledge, that of the experts. Characteristics of the subjective knowledge are limitations in space and time and frequent changes. By using CBR, this subjective knowledge can be described and represented by cases, and changes in this knowledge yield updated cases covering the newly obtained knowledge.

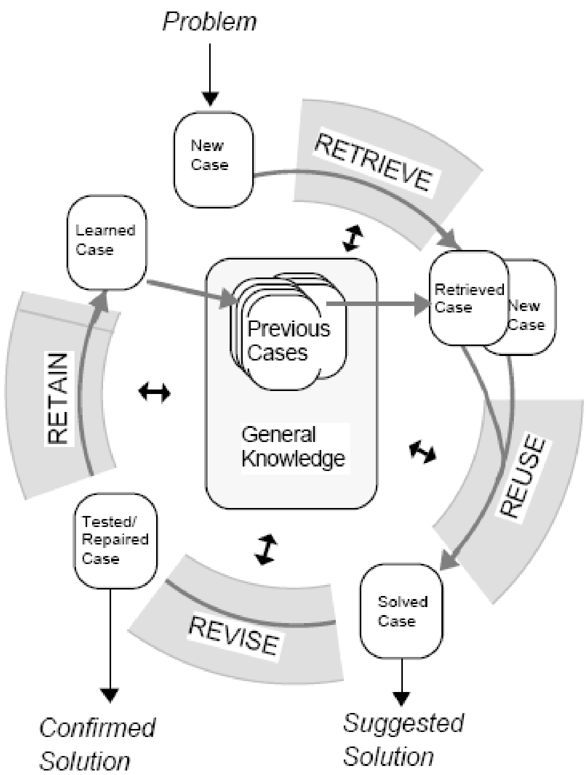


Figure 4.2: The four stages of CBR illustrated as a circle. Adapted from [39].

Aamodt identifies the main phases of CBR as a circle of four stages, see figure 4.2. The CBR cycle is activated by formulating the unsolved problem at hand as a new case and submitting it as input to the cycle. The four stages are explained in the following sections.

4.3.1.1 Retrieve

The new case submitted to the cycle is used to retrieve one or more cases from the case base of previously solved cases. The retrieval approach varies, as will be apparent in section 4.3.2. Some systems are based only on simple syntactic comparison of the case descriptors, while other use deeper semantic similarities during the retrieval phase. These systems are categorised accordingly as knowledge-poor and knowledge-intensive approaches as described above.

4.3.1.2 Reuse

During the reuse phase, the retrieved case will be combined with the initial case to form a new solved case. This includes finding the differences between the two cases and which parts that could be transferred from the solved case to the new case. The simplistic way of reusing is to consider the solution class of the solved case, abstract the differences and transfer the class as the solution class of the new case. This is only valid for small simplistic systems. The typical reuse process will have to take into account the differences of the cases and the fact that the solution cannot be transferred directly because of these differences. The solution of the new case needs to be an adaption of the old one with the differences of the case taken into account. Solution approaches include adapting either the solution of the solved case or the method used to construct this solution [43].

4.3.1.3 Revise

After constructing a new solved case in the reuse phase, the case needs to be evaluated to verify the validity of the solution. During the revise phase, the case and its solution will be evaluated by an instructor or applied to a real world problem. If the case solution fails, it will be repaired according to the properties of its failure. The main tasks of the revision is to learn whether the solution is successful or perform repairs if it fails. This evaluation often contains a great deal of delay before its successfulness is known. To take advantage of everything learned, it is important that the system always knows whether a solution has been successfully evaluated or not, or if the processing is still going on. It is equally important to ensure that solution failures are learned as well, to prevent the mistakes from happening again [38].

4.3.1.4 Retain

After completing the problem solving, the system needs to decide which parts of the new experience to keep. This includes decisions such as whether to store a new case in the case base, changing an already existing case, how to index the case for later efficient retrieval, and how to integrate the retained information into the already existing memory structures. If a new, and different solution method was used during the revise phase, a new case will be created and added to the case base. If not, an existing case might be generalised in such a way that it will include the information from the new case. After deciding what to keep and how to integrate it into the case base, the system is ready to use the newly acquired knowledge when retrieving matches for the next incoming case.

4.3.2 Applications of CBR

Applications using CBR has been developed in a number of domains throughout the last 25 years. To be able to understand the specifics of CBR, this section will elaborate three CBR systems, namely Protos (section 4.3.2.1), Casey (section 4.3.2.2), and Creek (section 4.3.2.3). Protos and Casey were developed for medical diagnosis, while Creek was originally developed for mud diagnosis in oil-well drilling [38]. Each of the applications are explained in the following sections.

4.3.2.1 Protos

The Protos system [44] by Porter/Bareiss is an approach to case-based classification in weak-theory domains and contains an alternative way to organise cases in the case memory. The domain in which Protos was originally used was auditory diseases. The original case of the CBR system is, in the Protos system, separated into cases and exemplars, with cases being new, unsolved instances and exemplars solved and retained cases. The exemplars are instances of a category, which is the representation of a real world concept. The membership of the exemplars to the category is described by assigning different degrees of importance to their features.

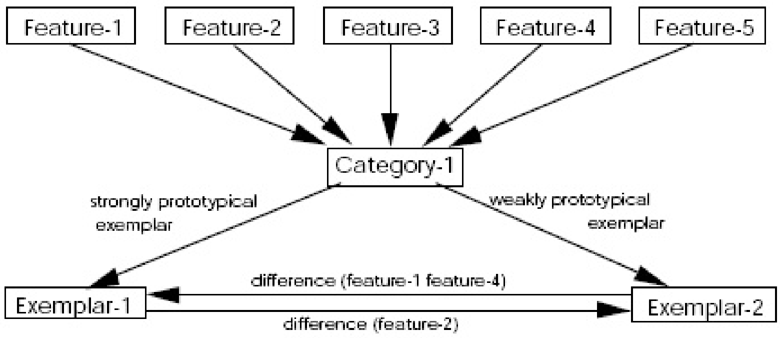


Figure 4.3: Index pointers of the Protos system between features, categories and cases. Adapted from [38].

The case memory of Protos is organised as categories, cases and index pointers [44]. Cases are associated with categories and the index pointers can point to either the case or the category. As mentioned in [38], there are three kinds of index pointers (see figure 4.3):

- Feature links (remindings) - pointers from a feature to a category or case
- Case links - pointers from a category to its cases
- Difference links - pointers between cases with common features and suitable discriminating features

The classification of Protos starts by collecting and combining the remains of the features of the new case. The most similar exemplar is then selected and the difference links from it is used to improve the match even more. The category of the selected exemplar is used to classify the new case into a specific category. The result is then suggested to an expert, which may reject it and provide an alternative explanation or accept it as it is. The evaluated solution is used during the revise phase to produce a verification. After being revised, the case will be added or already existing index strengths are updated. The learning of PROTOS is heavily based on user interaction to be able to explain the relevance of each feature and to verify the solution at all times.

4.3.2.2 Casey

Casey was developed for case-based diagnosis of heart failure diseases. As opposed to Protos, Casey was designed as a closed system with no user-interaction. It heavily depends on the causal model for its reasoning. To be able to reason without user-interaction, the cases of Casey contains large amounts of information. They contain not only the observed features, but also causal explanations of their diagnosis and a structure for storing the states of the heart failure model where evidence has been found in the cases. When retrieving a case from the case base these general causal states are used as a primary index and the observable symptoms are used as a secondary index. [39]. When reusing a previous case, the system attempts to copy the diagnosis of the retrieved case into the new case. In most cases this is not possible and an adaption of the solution is necessary. This adaption is based on modifying the structures of explanation that is contained in the retrieved case. If the case-based method cannot provide a solution, Casey uses a model-based method based on the causal model to find a solution. Learning from successfully solved cases includes updating the importance of features or storing the information as a new case along with its explanation [45].

4.3.2.3 Creek

Creek (Case-based Reasoning through Extensive and Explicit general Knowledge) was developed by Aamodt [38] as an architecture for knowledge-intensive case-based problem solving and learning, used in open and weak theory domains such as diagnosis problems in oil-well drilling and medicine. Problems are described by problem solving goals, solution constraints, and lists of findings. The solutions are typically a diagnosis and/or a certain repair that must be performed. Since Creek is a knowledge-intensive approach, its cases are enriched with explicit general domain knowledge and each case is viewed as a user experience.

The knowledge modeling approach in Creek is based on a combination of top-down, initial knowledge modeling, and bottom-up, knowledge maintenance. The tasks of the initial knowledge modeling are used to analyse the domain and problems at hand, and to design and develop necessary initial operations and conceptual models to be used by the system. The knowledge maintenance is activated when the initial modeling has ended and should ensure that the system knowledge are kept up to date and if possible refine it further. The combination of these approaches are illustrated in figure 4.4.

Knowledge models created in Creek can be viewed as semantic networks composed of nodes and edges, each explicitly defined in its own frame [37]. The nodes represent concepts and the edges represent relations between the concepts. The relation between two concepts is also what defines the concepts themselves and these relations are listed in the frame definition of the concepts in question. Figure 4.5 shows a semantic network composed by the three main knowledge types in Creek.

The explanation engine of Creek contains three steps of processing for each phase of the CBR cycle. The three steps are activate, explain and focus. The activation step makes sure that the

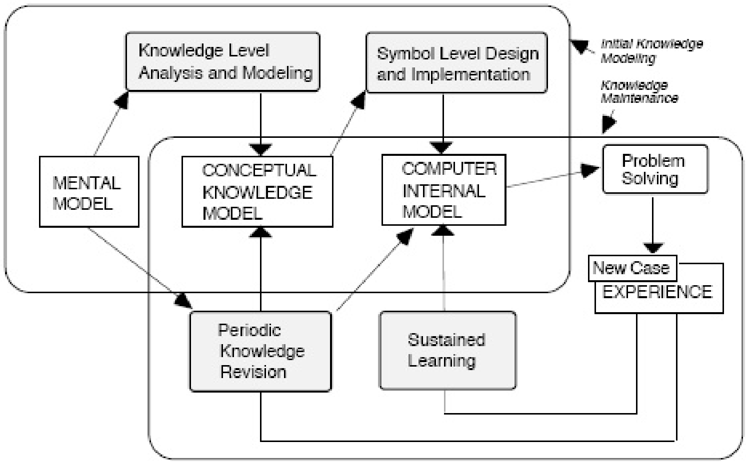


Figure 4.4: The cycle of knowledge modeling by combining a top-down and bottom-up approach. Adapted from [37].

relevant part of the semantic network is activated. The explain phase generates and explains information derived within the activated part of the semantic network, while focus tries selecting a solution that is valid according the goal of the task. Practically, during retrieval, a broad context are determined by spreading the activation phase through the network, followed by the index-based retrieval of matching cases and an explanation-driven selection of the best matching case [45]. After retrieving a matching case, the system attempts to copy the solution. If this is not possible, the system uses explanation-driven adaption of the solution by combining it with general domain knowledge. The solution (now possibly adapted) is evaluated by the user who gives feedback and the case status is stored for the case to be used during retrieval and reuse at a later stage. The knowledge learned is retained by trying to merge the cases in question, storing their findings, solutions (both successful and unsuccessful), and explanations. Case base maintenance is ensured by updating the strength of the indexes accordingly.

4.3.3 Strengths and weaknesses

Before choosing to use CBR as a tool for reasoning about knowledge in the medical domain, it is important to be aware of the strengths and weaknesses that must be considered. This section will describe some of the strengths and weaknesses inherent in CBR [45].

4.3.3.1 Strengths

CBR includes the ability to take advantage of historical knowledge without intensive knowledge acquisition from a human expert. This is easily done by obtaining cases from existing case structures, logs of errors and repairs, or other similar sources.

Assuming a reasonably representative case base, CBR offers shortcuts during the process of reasoning. If it is possible to find an appropriately matching case, the input case can be solved

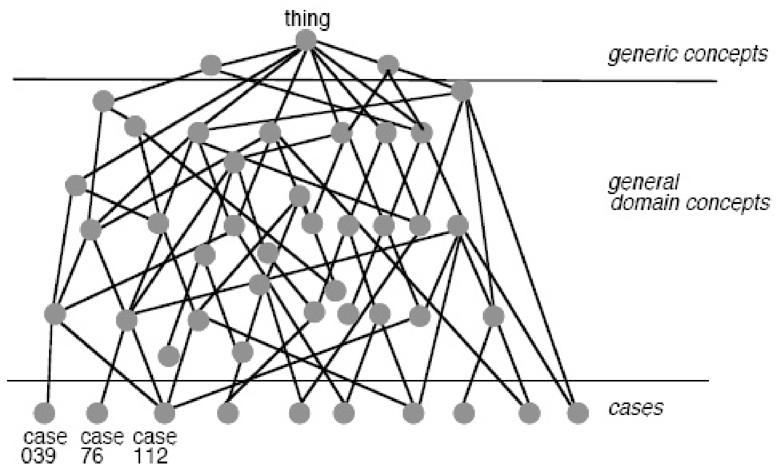


Figure 4.5: The integration of generic concepts, general domain concepts and cases. Adapted from [37].

more efficiently than by using rules and models.

A case-based reasoner using a suitable indexing strategy will make it possible to efficiently distinguish between target problems and finding an appropriate matching case. This strategy is a substantial part of the problem-solving power of CBR.

As opposed to rule-based systems, CBR does not require an extensive analysis of the domain knowledge. It is based on an additive model of knowledge acquisition which makes it possible to analyse domain knowledge as new cases are added or altered. The success of this additive model heavily depends on a good strategy for adapting case solutions, a meaningful representation of case knowledge, and a resilient structure of case indexing and retrieval.

4.3.3.2 Weaknesses

Using CBR may often result in a large case base restricting the trade-off between storing and computing. Without a good structure of the indexes, substantial overhead when storing and retrieving cases may occur. In a system of limited resources this overhead will further increase the overall computation time.

Cases of CBR may often not include any deeper knowledge of the domain in question. This enables situations where cases are applied to wrong situations because of the lack of explanation possibilities. A CBR system giving advices or providing solutions to domains such as the medical domain will suffer greatly and might cause wrong conclusions to be made.

The development of index structures and matching criteria requires carefully intervention of a human expert, potentially using vast amounts of time and resources. This will become a drawback if no standards can be declared and the task has to be carried out every time a CBR system must be tested. The shortcuts offered by CBR when finding a matching case come at the price of developing these criteria and indexing the cases efficiently.

4.3.4 Current problems

There are as of today no general CBR approach suitable for every possible domain with or without general domain knowledge. A majority of the current problems are related to efficiency and accuracy, which are crucial factors when considering CBR as a tool in a production environment. The next sections will describe some of the most common problems inherent in memory organisation of CBR and each of the reasoning phases.

4.3.4.1 Memory organisation

The main problem of memory organisation is the structural problem of choosing a case structure to ensure an efficient and powerful system for knowledge reasoning. After deciding on a specific case structure, the choice of index structure will further limit or aid the search and maintenance of cases. The choice of this index structure is not an easy task, and will further complicate the issue of memory organisation. It is important, when using a system based on general domain knowledge, to integrate these concepts without limiting the complete system.

4.3.4.2 Retrieval

When retrieving cases similar to the input case, the use of indexes must be considered. The retrieval is always present and an efficient search method based on these indexes needs to be developed. To be able to assign a degree of similarity to different cases, the relevance of each feature needs to be decided upon. This depends on the domain and the problems the system is expected to solve, and must be adjusted accordingly. After comparing cases and assessing their similarities, problems related to the use of general domain knowledge and the use of previous cases must be handled. Their integration into the retrieval phase is necessary to obtain subtle reasoning processes.

4.3.4.3 Reuse

The successful retrieval of a similar case leads on to the reuse phase. The main problems to be considered here are the transfer and adaption of the existing solution and solution method onto the new, unsolved case. If the input case matches a known case, how should its solution be transferred to the new case efficiently and without loosing knowledge? If the cases are not a complete match, the known solution must be adapted to the new case. What is an efficient way of doing this, and how should it be decided which parts that should be ignored?

4.3.4.4 The learning phases (revise and retain)

The most complex phases of the CBR cycle are the revise and retain phases, often combined into a learning phase. A variety of problem areas evolve from the task of learning from recently acquired information. An approach to extract features from the solution at hand is needed and questions concerning whether to separate the cases or to split them up must be answered. To develop an efficient learning phase, the learning of indexes must be considered as well as the ability of the system to generalise several instances of specific knowledge into a more general form. A CBR system will not only have to learn, but it should also have the ability to forget. To avoid complex and unmanageable case structures or storing large amounts of specific knowledge instead of a generalised one, a way of forgetting knowledge would be appreciated during case base maintenance.

4.4 Neural networks

This section will first describe the principles of neural networks, before explaining the initialisation and training phases. In order to compare neural networks and case-based reasoning, its strengths and weaknesses are presented in section 4.4.2.

4.4.1 The principles

The main idea of neural networks is to model the behaviour of a brain when presented with a specific task. The brain is made up of a large group of interconnected neurons. These neurons receive an input stimuli, and transmit a new stimuli to other neurons based on the strength of the input stimuli. The neurons perform a weighting of their input to control the amount of stimuli transmitted to another neuron. The man-made neural networks try to simulate this behaviour using artificial neurons. By first weighting the input given to the neuron, and then applying a specific function (for instance a sigmoid function) to the result, the size of the output signal is determined. The objective is for the network to produce a single high output at one output node for each class of input problems. By making sure this criterion is fulfilled, the network is able to recognise and separate the different input problems it is given.

4.4.1.1 Initialisation and training

The neural network is organised with an input layer, an output layer, and a number of optional so-called hidden layers in between, see figure 4.6 for an illustration of a network with a single hidden layer. The neurons at each of these layers are interconnected to the neurons at the next level, allowing information to flow from the input layer, through the hidden layers, and produce a result at the output layer. Since the input layer consists of a number of neurons, the problem to be solved must also be transformed into a shape which is suitable as input to the network. A common form of representing a problem is by transforming it into a vector consisting of nodes where each node correspond to a neuron in the input layer. When the vector is first presented to the network, weights are initialised at the links between the neurons. These weights will, as mentioned above, control the output of each node in the neural network. After propagating the result through to the output layer, the weights must be changed so the network can identify this particular input (by having a single response at the output layer). The main method for doing this is by back-propagating the result through the network. The weights are then changed so that only one node at the output layer is activated when the network is presented with a particular vector. The user will then know that if the network is given an unknown input vector and this triggers the particular node at the output layer, the network has recognised the input vector.

If the network is only trained with one version of each vector, it is unable to recognise variations of the same vector. This is a typical problem in image recognition problems, where

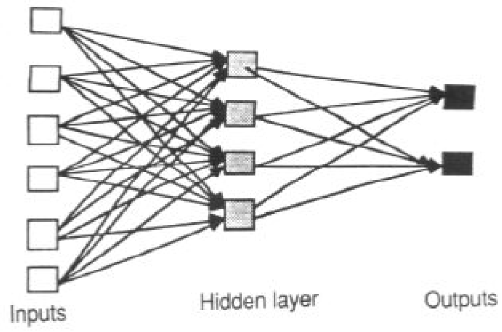


Figure 4.6: A simple neural network with a single hidden layer.

the idea is to recognise variations of a defined standard. To cope with these problems, the network must be trained. After initialising the network, as mentioned above, the network is given several noisy versions of each input vector. This enables the network to also recognise these shapes, and hence it can recognise arbitrary versions of the input vector. When performing this training step, it is important not to overtrain the network. Minimising the error during training does not guarantee that the network will recognise more general shapes of the object at hand [46].

4.4.2 Strengths and weaknesses

This section will discuss the strengths and weaknesses of using neural networks in the field of image processing, and particularly in image analysis. The findings of this section will be used when deciding which method to use in section 4.6.

Although neural networks are heavily used in several image processing tasks, including image analysis/understanding, their most common use is in the lower levels of the image processing pipeline, i.e., segmentation and recognition. Egmont-Petersen et al. [46] say it quite well, by stating: *We feel that image understanding is the most dubious application of ANNs[1] in the image processing chain*. Neural networks are not particularly suited for higher level processing tasks as they rely primarily on pixel intensities as input, and it is difficult to include for instance a priori knowledge in the network. According to Egmont-Petersen et al. [46], more than 60 % of the neural networks they have examined, use only pixel intensities as input to the neural network. This implies a large number of input nodes when working with only modest-size images. If additional information is added, this will further increase the complexity of the network. Since one of the prerequisites of a image analysis module is the inclusion of a domain knowledge model, a neural network approach will largely suffer from its inability to obtain and use the information present in this model.

In addition to the large number of input nodes, it is also difficult to specify the architecture of the network itself. It is not known prior to testing how the number of nodes and hidden layers will affect the result. To be able to design a working network, a combination of best practice and careful training should be applied.

[1] Artificial Neural Networks

Another significant drawback when using neural networks is their inability to analyse a problem which they have never seen before. If the network has not been trained with a specific input vector, and a variant of this vector is given to the network at a later stage, it will not be able to analyse the information present in the vector. The network will probably converge to a solution, but this solution does not have to be the best model of the data. This result implies that the network must be trained whenever a new class of objects should be recognised. This will severely limit its applicability, as the network cannot be trained to handle every possible situation. A better approach would be to process the unknown input vector by consulting the domain knowledge model before presenting a preliminary result to the user. It should then be possible for the user to actually infer how this conclusion was made in order to check the correctness of the analysis tool.

This leads to the probably the most serious limitation of neural networks, their black-box behaviour. Because of the internal couplings and often large number of neurons in each layer, one can only give the network an input vector and obtain the result from the output layer. There is no possibility to understand or interpret why the network gave the specific result. In a medical domain, this is a severe disadvantage. If a certain abnormality is found it must be possible to understand the reasons for why the specific diagnosis is made.

The main advantage of neural networks is the fact that they perform well at several different problem areas. Their versatility allows them to adapt to the problem at hand. This is often important, since it can then be used at different stages of the image processing pipeline with only minor modifications. This is, however, not very relevant to this thesis.

As mentioned above, the only work needed by the user after training the network is to give the network its input, and the result is immediately available. This is often problematic, because of the black-box problem, but in some cases it can be beneficial. This feature makes the networks easy to use. If one can be certain that the network performs according to its specifications, a user without any previous training can use them to get reliable results.

Some of the limitations of the neural network can be reduced using neuro-fuzzy networks [46] which are neural networks based on fuzzy logic. Fuzzy logic introduces the concept of multivalent sets, which means that an item can be a partial member of one or more sets. Compared to the bivalent sets of standard set theory, this makes it possible to model continuous systems correctly. It is however important to understand that the degrees of fuzziness is not the same as probability percentages [47]. Probability percentages is a measurement of whether an incident will occur or not, while fuzziness is the measurement of the degree to which some condition is fulfilled or something occurs.

Combining neural networks and fuzzy logic into neuro-fuzzy networks makes it possible to incorporate the uncertainty often present in complex domains into the training and learning processes of a neural network. Domain experts are able to represent domain knowledge as rules based on fuzzy logic. The parameterised nature of neural classifiers makes it otherwise hard to develop constraints using previous knowledge.

4.5 Domain relevance

To able to choose one of the two previously described methods, it is important to recognise the specific challenges present in the working domain. The medical domain is complex and demands a number of restrictions to be carefully considered. This section will present the most important ones and describe how the two methods proposed attempt to solve them.

- The medical domain requires that general domain knowledge is considered before making a diagnosis.
  - A knowledge-intensive CBR system will by default require the use of general domain knowledge as part of the cases. This knowledge enables the system to reason, not only syntactically, but also semantically, and pragmatically.
  - A standard neural network does not take general domain knowledge into account except from those parts contained in the input vector. Fuzzy networks can increase the amount of general domain knowledge by specifying a set of rules[2].
- The complexity of the medical domain requires that previous knowledge is used when solving a present problem.
  - A CBR system uses an additive model of knowledge acquisition, making it easy to incrementally add the knowledge learned from each case to the case base and models are then used to reason about the present problem.
  - Training a neural network implies feeding it with previous knowledge. When solving a given problem presented by an input vector, this knowledge is used as the basis for reasoning. By adjusting the membership (activation) functions in a fuzzy network, the knowledge given in the input vectors during training can be further specified.
- The medical domain requires the ability to reason about complex relations, as a diagnosis can rarely be made from a single observation.
  - The reasoning of CBR can be supported by causal models of the domain in addition to the general domain knowledge included in the case base. This renders complex reasoning possible, taking a range of possibilities into account. Causal connections and other relations are included when new knowledge is learned, making it available to the next problem solving session.
  - The ability to handle complex relations with neural networks depends on characteristics such as the number of layers and the type of input vectors. Using a fuzzy network will further increase the robustness of the system and make interpretation easier.
- The complexity and consequences of a diagnosis requires that the reasoning can be verified to understand how the system arrived at the specific conclusion.
  - A knowledge intensive CBR system, including general domain knowledge, will provide explanations for the matching of cases and the adaption of earlier solutions to new cases. These explanations are retained in the system when the solution is confirmed and can be inspected by the user to ensure the justness of the particular diagnosis.

[2]On the form IF X AND Y THEN Z.

– Neural networks are often considered black-box systems because of their complexity in the number of layers and couplings between the layers. This makes it very hard to provide an explanation for the result at the output layer based on the input vector.

4.6 Choice of method

The previous section listed the most important restrictions introduced by the medical domain when creating a system within this specific domain. Two methods, case-based reasoning and neural networks, were inspected to see how well they would handle these restrictions.

The main common factor when trying to make a diagnosis is the need to consider knowledge. This could be previous knowledge of similar events or discoveries, or general knowledge that applies to the medical domain. To have a system automatically make a diagnosis means that it has to incorporate these sources of knowledge into its decision core. Some functionality to read, process, and store knowledge must be present. By utilising knowledge, the system will be able to take advantage of complex relations for reasoning, which is hard using only syntactical information. In addition, the system must have the ability to learn from previous experiences in order to avoid making the same mistakes twice and to ensure a greater rate of success. Both CBR and neural networks (fuzzy) have the ability to exploit both general domain knowledge and previous knowledge within their computations. CBR, and especially Creek, incorporate general domain knowledge as an integrated part of the cases using a semantic network of nodes and edges. This ensures that the integration of general domain knowledge is already a built-in feature. Neural networks do not have the same opportunity as the general domain knowledge must be presented to the system by experts using for instance rule definitions.

Another important property needed for automatic diagnosis is the ability to check and verify the conclusion reported by the system. This is of outmost importance when mortality or other serious consequences are expected. The reasoning of neural networks is often considered to be 'black-box', making it very hard to verify that the answer given at the output node is correct according to the input vector. This significantly weakens the usefulness of neural networks for automatically making a diagnosis, as a wrong conclusion can have critical consequences. CBR systems on the other hand, such as for instance Creek, use predefined structures to organise knowledge and information. Using a semantic network of nodes and edges makes it very easy to follow the relations (edges) in order to understand and verify the reasoning of the system within a knowledge model.

Based on the clarifications given in the previous paragraphs, CBR is chosen as the method of reasoning used in this thesis, and in particular Creek is chosen as the architecture. Creek was designed for open and weak-theory domains, making it suitable for usage in the medical domain. It integrates problem solving and learning into a single combined architecture. Practically, this thesis will use the jCreek tool for case matching. JCreek will be described in the next chapter.

Chapter 5
System solution

The three previous chapters have elaborated the subtasks presented in chapter 1. A theoretical background to the concepts of standard view planes, feature extraction and understanding, and feature analysis has been provided. Each chapter concluded on how to solve the challenges related to each of these concepts. This chapter summarises the chapters by identifying the components required to solve the tasks of this thesis. A practical approach will be presented, in contrast to the theoretical approach previously given.

This chapter is divided into three parts. Section 5.1 summarises chapters 2, 3, and 4 by identifying a list of challenges and solutions that must be incorporated in a working prototype. Each of these components, constituting the prototype, are explained in section 5.2, while section 5.3 explains the workflow between them.

5.1 Task elaboration

The previous chapters have, as mentioned above, provided the theoretical background for solving the tasks of this thesis. This section presents the challenges to be solved, and the solutions adopted. In addition, the transitions from data to information, and from information to knowledge, as presented in section 4.2, are presented.

The first challenge requiring a solution is the extraction of regions from the image. Section 3.3.5 presented the segmentation routine chosen by this thesis. A generic segmentation routine uses only the intensity of the pixels as input, and no interpretation of the data is performed. This implies that the processing is performed at the data level when referring to the classification of section 4.2. When applying the live-wire approach, one can argue that the output is information. This is based on the interpretation of the image contents by the medical expert guiding the segmentation routine. After obtaining the information about the regions in the image, they must be described in order to discover which organs they represent. Section 3.5.3 presented several region-based descriptors that will be applied to describe each region. The information is further elaborated during this process, but still remaining at the information level. In order to reach the knowledge level, a matching procedure must be executed. This procedure was previously explained in section 3.6. The result is knowledge about which organs the extracted regions most likely represent.

The matching procedure is the last one of the image processing pipeline. The obtainment of knowledge about the organs (and possible symptoms of abnormalities) makes it possible to reason about these findings. Reasoning requires that knowledge about fetal anatomy and abnormalities, as presented in chapter 2, is stored in a domain knowledge model. A common representation of both the obtained knowledge and the stored knowledge is needed to make reasoning possible.

The last component required is the analysis tool. Section 4.3 presented the CBR approach that will be used by this thesis. This enables reasoning about the contents of the extracted standard view plane, and a diagnosis can be made based on existing knowledge about fetal anatomy and congenital abnormalities.

This section has identified the components needed to make a prototype that can automatically make a diagnosis based on the findings of standard view planes. The next section will further explain these components.

5.2 Components

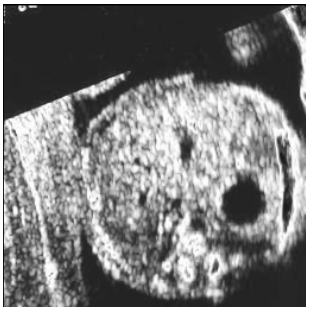
The previous section presented a further elaboration of the main task, and identified the challenges that had to be handled by the application to fulfill the main task. This section focuses on a further decomposition of the superior solutions presented and examples from the developed prototype are included as a proof-of-concept. Section 5.2.1 will further elaborate on the object extraction. Object description on a component level is described in section 5.2.2, while the details of the object recognition is discussed in section 5.2.3. Section 5.2.4 will explain how the existing knowledge of the medical domain is organised as a domain knowledge model, and section 5.2.5 describes the reasoning within this model.

5.2.1 Region extraction

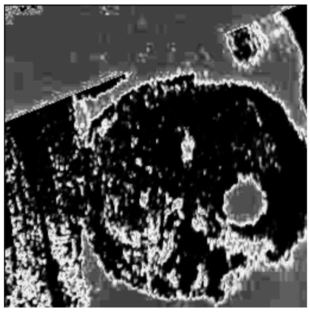
The task of extracting objects often requires several pre-processing steps to remove noise and artifacts from the original image. The most common forms of noise corruption present in ultrasound images have been presented in section 3.1, while section 3.2 presented properties that had to be present in a segmentation routine applied to ultrasound images. The properties presented were directed towards the segmentation step, but by pre-processing the images before segmenting them, one could further improve the quality of the objects extracted. Common examples of pre-processing used to increase the quality of the image are histogram equalisation, image smoothing, image sharpening, median filtering, and edge detection [36].

The method applied to increase the segmentation quality when using the live-wire method, is to train the image by adjusting the cost matrix (see appendix A.1). First, a rough outline of the object contour is made. This adjusts the initial cost matrix to make sure the borders close to the outlined object contour is given increased importance compared to edges further away from the object. This implies that when trying to segment the object a second time, the contour is more likely to follow the contour of the object to be extracted. By performing several training iterations, and hence adjusting the cost matrix several times, an overall increase of the quality of the object segmentation is achieved. The number of training iterations needed depends on the strength of the original object boundary. If it is sufficiently strong, the outlined border is attracted to it, producing a suitable region border. If the object is corrupted by noise, the training iterations become more important in order to segment the object correctly.

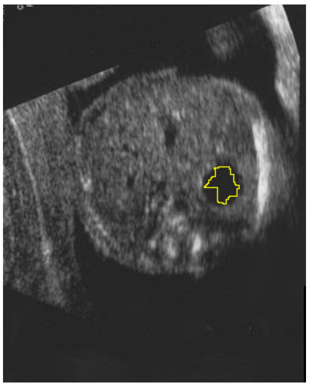
The concept of the training iterations is illustrated in figure 5.1. Figure 5.1(b) shows the image representing the cost matrix of the original image after training it for three iterations. It is apparent in this image, as bright colors indicate strong edges, that a contour made within



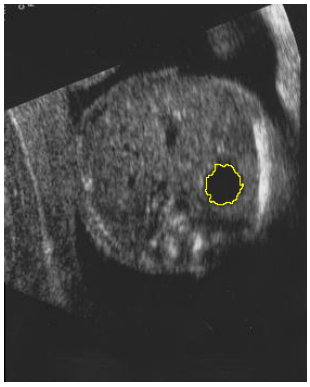
(a) Result without training



(b) Result after three iterations of training



(c) The resulting border region without training



(d) The resulting border region after three iterations of training

Figure 5.1: Illustration of segmentation using the live-wire method before and after training.

the region of the stomach bubble is pulled out towards the true edge. When comparing images 5.1(a) and 5.1(b), it is evident that the border of the stomach bubble is much more distinct in the trained version than in the untrained version. This is also reflected in the segmented images 5.1(c) and 5.1(d), where the region boundary is much closer to the true boundary in figure 5.1(d) than in figure 5.1(c).

5.2.2 Region description

The task of region description was previously elaborated in section 3.5, which concluded that region-based descriptors were the best choice for describing each region found in a standard view plane. This section present the feature vector which is later applied during object recognition.

The feature vector consists of the four descriptors[1] presented in section 3.5.3. A complete feature vector describing the regions in figure 3.7(a) is found in table 5.2.2.

As can be seen from the table, there are several statistical moments which all appear to have the value zero. This is because a limit has been set as to how small a moment can be before

[1]The statistical moments consist of seven invariant moments.

	mad	rib_right	umb_vein	stomach_bubble	rib_left	spine
Area	0.4168	0.0147	0.0028	0.0169	0.0159	0.0151
Circumference	0.0127	0.0032	0.0009	0.0025	0.0033	0.0030
Compactness	26.7033	47.0069	19.713	24.8923	48.1994	41.6400
Moment 1	0.1859	0.5908	0.2060	0.1654	0.4756	0.2559
Moment 2	0	0.3169	0.0129	0.0006	0.1664	0.0012
Moment 3	0	0.1255	0.0007	0	0.0077	0.0058
Moment 4	0	0.0138	0	0	0.0041	0.0029
Moment 5	0	0	0	0	0	0
Moment 6	0	0.0078	0	0	0.0005	0
Moment 7	0	0	0	0	0	0

Table 5.1: A feature vector describing the regions of figure 3.7(a).

being discarded. The reason for this thresholding, is that a very low number at one of the features can easily result in a large deviation when performing object recognition in the next step of the pipeline.

5.2.3 Object recognition

The task of object recognition was previously described in section 3.6 where the procedure of matching unknown region descriptors and descriptions of known organs was presented. This section will further elaborate on how the knowledge of known organs should be stored and organised in order to utilise it efficiently. An alternative to the direct matching between regions and organs is presented in section 5.2.3.1.

To store knowledge about known organs, an organ database is created to keep records of each known organ. Each record contains similar feature vectors, as presented in the previous section, to enable an efficient matching procedure. To utilise such a database, there are several properties that must be fulfilled. First of all, the database must contain enough records to correctly identify the most common organs present in each standard view plane. The description of each organ should also incorporate the natural variations of each organ. The natural variation is often described statistically using nomograms showing for instance estimated size of an organ at certain gestational ages, see for instance section 2.2.2.3, where the size of several important features of the brain are described.

As the application aims at finding abnormalities in the standard view plane, the database must also contain feature vectors describing symptoms and features related to abnormalities. When describing the features related to abnormalities, the natural variation is not always known, causing a more limited description to be stored in the database. As several occurrences of the same feature is found, the record in the database is updated by incorporating the newly obtained knowledge to create a more detailed description of the variation of the feature. The database is also updated whenever organs are marked as unknown, as stated in section 3.6.

To enable a more efficient matching procedure, the organs described in the database are categorised as belonging to a particular standard view plane. This ensures that when finding the best match of a particular organ, only organs of the same standard view plane are considered.

5.2.3.1 Using CBR for object recognition

An alternative to using direct descriptor matching when trying to recognise the objects present in the standard view plane, is to incorporate a CBR system at this level. The advantages and disadvantages of applying CBR for object recognition will be discussed before a specific system is presented.

The main advantage of using a CBR system is the ability to apply previous solutions to similar cases as the one in question. When presented with an unknown organ, it will try to adapt an existing solution to the problem, or external expertise is consulted to solve the case. The case is then added to the case base for later retrieval. The case base itself is also one of the advantages of a CBR system. Because CBR will be used during analysis, parts of the same case base can be reused when doing object recognition. The output of this module can be directly applied at the next step, without much need of changing the results.

Another advantage is the possibility to easily include more information about the contents of the image when performing the object recognition. The approach suggested in this thesis is only to compute the degree of match between two organs. By adding information indicating certain organs, a more reliable matching procedure could be developed. This is easier to incorporate in a CBR system, as it is designed to match cases and the contents of the cases are not defined. The approach of this thesis is not as dynamic as a full scale CBR system, and will suffer from the lack of additional information.

The main disadvantage of using CBR in object recognition, is its increased complexity compared to a simpler matching procedure. This is the main reason why CBR is not used in this thesis. The descriptor matching gives a sufficient recognition rate, and by introducing CBR the recognition rate will not increase sufficiently to justify the increased complexity of the system.

Francois and Medioni [48] state that a CBR system should be applied if the reasoning should be based on actual previous experiences. By using the organ database, which represent the natural variation of each organ, previous experience is already modeled, without having to introduce a more complex system.

Bendiksen [49] presents a method for using CBR during object recognition. In this work, each object is described using descriptors, but additional information about the fetus is added to the case. The increased amount of information enables a more reliable discovery of certain fetal organs. The most important addition is to include knowledge about abnormalities or deficiencies affecting the fetus. It should be obvious that knowledge about abnormalities is a significant advantage when trying to recognise objects. Abnormalities imply that certain organs should be visible, and hence one can narrow the search space of possible organs. The problem of this approach, is that this information is not present in our situation. As part of the task is to make a diagnosis, no knowledge about the abnormalities is present at this stage.

5.2.4 Domain knowledge organisation

The knowledge obtained from analysing the standard view planes of a fetus is not sufficient to detect abnormalities or make a diagnosis. General knowledge about the domain is also needed to be able to understand and reason within the relations and connections found in the standard view planes. This section will consider this general domain knowledge accompanied by the image knowledge acquired by the components of section 5.2, and try to organise it for further analysis. It will also consider the tools used in this thesis to perform the organisation.

The medical domain is known to be a open and weak theory domain [38]. This means that no established theory exists to explain every detail of this domain. To enable reasoning and understanding in such domains, an extensive body of general domain knowledge is required. This body should include relevant domain concepts, relationships, and constraints modeled as a tightly coupled structure with different levels, depth and details [38].

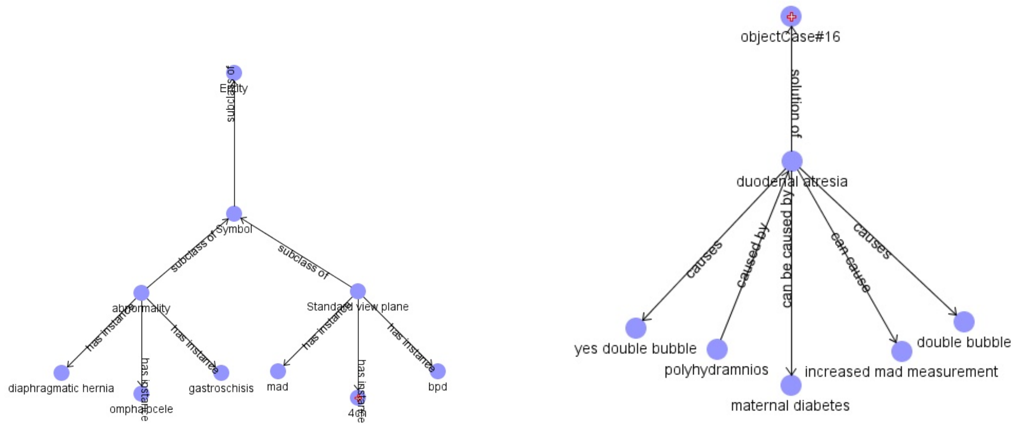
It is important to be aware that even though the medical domain is a weak domain, there is no lack of general knowledge available. One should rather consider the fact that the knowledge is more or less certain or stronger or weaker, instead of just true or false. Without a collection of ground truths, it is difficult to apply methods of inference and reasoning. A possible way to compensate for the lack of strong knowledge, is to include more general domain knowledge in a coherent knowledge model.

A medical domain knowledge model has to contain a complete set of knowledge of the fetal anatomy and a sophisticated collection of the most common abnormalities expected to be found during a routine ultrasonography. Because of the size of the medical domain, one can expect that only a subset of every possible abnormality can be covered in an initial knowledge model. This makes it even more important to include general domain knowledge into this model. Aspects such as cooccurrence of abnormalities and relations between symptoms and abnormalities must be represented in the knowledge domain model to be able to make a diagnosis.

5.2.4.1 Knowledge acquisition

The first step of creating an initial domain knowledge model is the process of getting hold of the current domain knowledge. This acquisition of medical domain knowledge depends heavily on the interaction with human medical experts. They carry massive amounts of subjective knowledge that must be extracted, understood, and organised as general domain knowledge. Even though humans are good at solving problems within a specific domain, it is not always easy to explain how they are solved and describe the knowledge that is being used. Different approaches exist to consider the creation of an initial domain knowledge model. Schreiber and Wielinga [50] distinguish three different processes: knowledge identification, knowledge specification and knowledge refinement. Aamodt [38] divides the problem of knowledge acquisition into initial knowledge modelling and knowledge maintenance. The processes of knowledge identification and specification can be considered inherent parts of the initial knowledge modeling from Aamodt. The goal of this first step is to capture the competence needed and to provide a knowledge environment for problem solving.

The results of the knowledge acquisition of this thesis was previously provided in chapter 2. As mentioned, the target of this thesis is not to provide a complete set of general domain knowledge, but rather to present an initial approach of a possible solution. The general domain knowledge of chapter 2 was presented within the surroundings of the standard view planes chosen for further analysis, and should provide an example as to what knowledge would be preferable. The landmarks, descriptors and abnormalities are all important parts for solving the main task of this thesis.



(a) An example of the structural dimension.

(b) An example of the causal dimension.

Figure 5.2: Dimensions of the domain knowledge model.

Apart from the purely medical symptoms described in section 2.5, it is important to note the symptoms based on causalities, such as for instance diabetes, smoking, and drugs. They provide equally important causes for the search of abnormalities as the anatomical ones. This will be further elaborated in section 5.2.4.2.

The acquisition of general domain knowledge will potentially produce vast amounts of unstructured knowledge applicable to the problem(s) at hand. To take advantage of this knowledge, it is important to develop a suitable representation which enables efficient and logical modelling. The next section will consider the representation approach applied in this thesis.

5.2.4.2 Knowledge representation

A deep knowledge model should describe both the concept definitions and basic principles which the operational knowledge [38] is based on. Due to this complexity it should contain several dimensions of representation. The initial domain knowledge model presented in this thesis uses a structural and causal dimension.

The structural dimension utilises inheritance to ensure that a strict hierarchy is applied during retrieval, reasoning and learning. The causal dimension specifies the causes of the effects, the symptoms of the abnormalities in our case. Examples of both the structural and causal dimension are given in figure 5.2(a) and figure 5.2(b). Standard view planes of fetal anatomy and the general class of abnormalities are both reflected as subclasses of the general class of symbols. They both have more concise instances specified in the model. The example from the causal model shows an abnormality called duodenal atresia (2.5.2.3). Causal relations such as 'causes' and 'can cause', describe the causal relationship between the abnormality and its symptoms.

A knowledge domain model is represented by a semantic network of nodes and edges. The nodes of the network represent concepts of the domain, and the edges represent relations between these concepts. Each node (i.e., concept) is defined in a separate frame. Each frame contains slots, representing the properties of the specified concept. The properties indicate

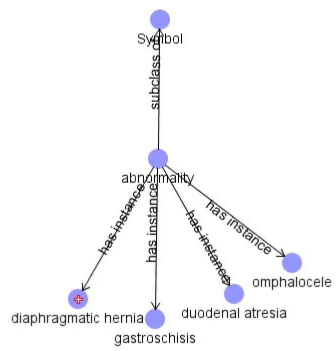


Figure 5.3: A snapshot from TrollCreek showing a few concepts and their relations.

the relations between the concepts. The slots are further divided into facets which describe the value type and value contents of each slot. An example illustrating these terms is given in figure 5.3. The figure shows a general class of concepts called 'Symbol' and its structural relationship with the specified concept of abnormalities using the 'subclass-of' relation. The class of abnormalities is further specialised as instances of this class with the 'instance-of' relation. Every node shown is a separate frame, and the 'subclass-of' and 'instance-of' relations are slots in the specific frames. The facets of the slots are value-types simply expressed as 'value'. The value-contents of each slot is the concept which it is linked to, by the given relation. A possible value-contents of 'abnormality' is 'omphalocele' within the slot 'instance-of'.

The knowledge base is separated into two kinds of knowledge. These are general domain knowledge, represented as concept classes and relations, and specialised domain knowledge, which are instances of concepts and previously solved cases. The general domain knowledge is exemplified by figure 5.3 with the 'Symbol' and 'abnormality' classes, and the specialised domain knowledge is shown as the instances 'diaphragmatic hernia', 'gastroschisis', 'duodenal atresia' and 'omphalocele'.

Cases, both solved and unsolved, are represented as concepts within the mentioned semantic web of concepts and relations. Each case inherits its superclass, 'Case', and contains a number of findings related to it. Solved cases contain solutions and possible explanations as well. The cases are integrated into the semantic network by having their findings defined as concepts in the same network. An example from the initial domain knowledge model is given in figure 5.4.

When a case is solved and queued for permanent storage in the case base, each finding is assigned a relevance to its solution. This is done automatically or by asking the user. The relevance of each feature is separated into importance and predictive strength. The importance of a feature describes how important the existence of this specific finding is to the classification of a case. Importance is assigned as irrelevant, informative, characteristic, or necessary, with characteristic as the default value. The predictive strength of a feature indicates how important this feature is in predicting one particular solution to the case. Predictive strength is assigned as spurious, indicative, strongly-indicative, or sufficient, with strongly-indicative as the default value.

The majority of unsolved cases created based on the knowledge extracted from the standard

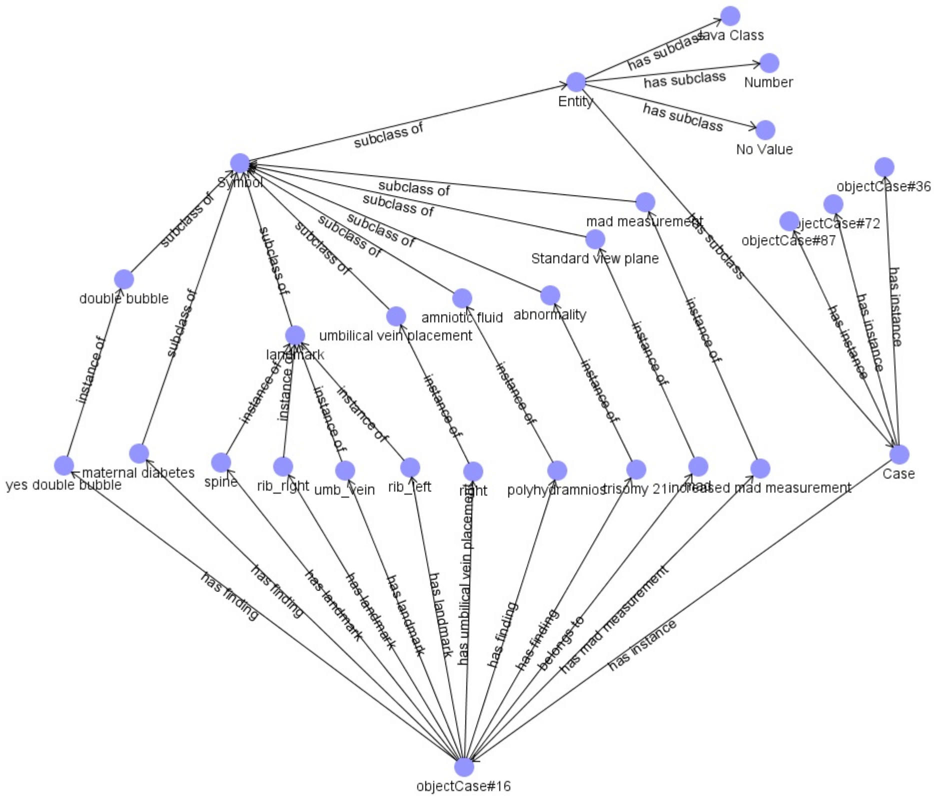


Figure 5.4: The integration of cases into the semantic network of general domain knowledge.

view planes, are confirmations of a perfectly healthy fetus. To minimize the amount of case processing, a procedure for coping with these cases has been suggested. The distinction of healthy cases is based on the fact that these cases should only contain findings confirming their standard view plane belonging. No findings representing symptoms of abnormalities should be present. This classification of findings makes it possible to execute a filtering of healthy cases before any case processing is performed.

5.2.4.3 Tools

As mentioned in section 4.6 this thesis uses jCreek [51], a Java implementation of the Creek architecture. JCreek implements the CBR cycle of section 4.3.2.3 and the explanation engine described in section 4.3.2.3. In jCreek, the knowledge model is modeled as a semantic network of nodes and arcs. Each node equals a concept in the domain and the arcs represent relationships between these concepts. The graph-like representation of knowledge with nodes and arcs offers the possibility to present the model as a graph. This thesis uses the TrollCreek [51] knowledge editor for this task and examples are shown in figures 5.2(a), 5.2(b), and 5.3. It is important to note that TrollCreek only focuses on the retrieve phase, and reuse at an early stage and is therefore not a complete tool. One important issue favouring TrollCreek, however, is its ability to perform case matching at any time during system development.

Listing 5.1: The adapted JCXML structure used to create new cases from knowledge extracted using image processing.

```
<?xml version="1.0" encoding="utf-8"?>
<root xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" ...>
   <case description="..." name="...">
      <finding relationType="belongs to" toEntity="MAD"/>
      <finding relationType="has umbilical vein placement" toEntity="
         right"/>
      <finding relationType="has landmark" toEntity="rib_right"/>
      <finding relationType="has landmark" toEntity="rib_left"/>
      <finding relationType="has landmark" toEntity="umb_vein"/>
      <finding relationType="has landmark" toEntity="stomach_bubble"/>
      <finding relationType="has landmark" toEntity="spine"/>
      <finding relationType="has finding" toEntity="malformed abdomen"/>
      <finding relationType="has finding" toEntity="polyhydramnios"/>
   </case>
</root>
```

With a means of acquiring and representing the general and specific domain knowledge at hand, a model containing both of them can be created. The specified knowledge of an image processing system is by nature different from the general domain knowledge created within a knowledge network editor. To cope with these differences, a common form of representation has been suggested using XML. This enables the system to transform the specific case knowledge of image processing as new case input to the knowledge network editor. An adaption of the JCXML from jCreek is used, and an example is presented in listing 5.2.4.3. As can be seen from this listing, the findings presented match the anatomical landmarks and descriptors, as well as symptoms of abnormalities presented in chapter 2.

5.2.5 Knowledge analysis and reasoning

The output of the previous section was an initial domain knowledge model. This model is a first attempt to create a representation of the general domain knowledge paired with the specified knowledge of the images analysed and should be treated accordingly. It cannot be considered the medical truth nor the absolute representation of the abnormalities as hand, but rather as part of a prototype.

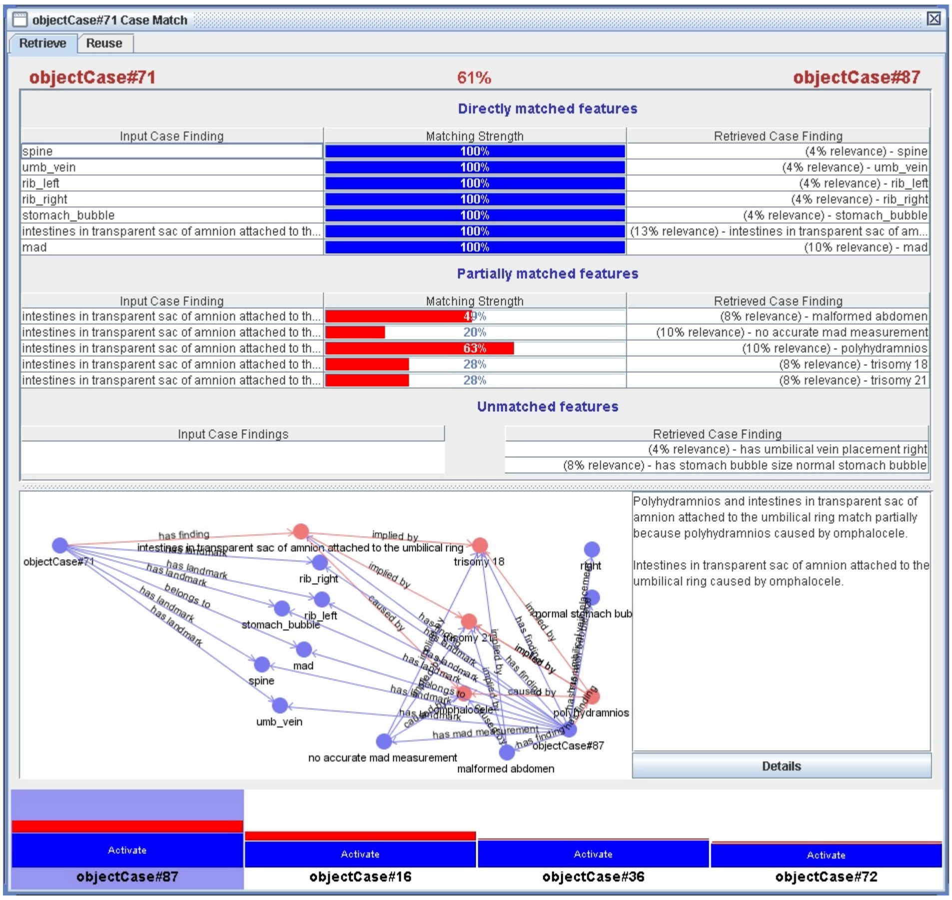


Figure 5.5: The matching of a new unsolved case against solved cases from the case base.

The acquisition and representation of knowledge were the first two steps required to be able to analyse standard view planes of fetal anatomy. This section will describe the next step, the actual analysis, based on the initial domain knowledge model developed and the properties of CBR and Creek already described in section 4.3.1.

The main target of the analysis is to understand or explain the consequences of what is observed in a standard view plane during an ultrasound examination. The combination of the previously described components will produce a model of the general domain knowledge and instances of observed knowledge. The remaining task is to reason within this knowledge and try to extract consequences and possible diagnosis. A required condition of such a procedure is an established base of experiences (cases), which constitutes the foundation for similarity assessment. The initial knowledge model contains a limited number of cases describing the set of symptoms required to discover the abnormalities presented in chapter 2.5. An example of such a case was previously illustrated in figure 5.4.

Given the complexity of the task, this thesis has focused on the establishment of an initial domain knowledge model and the application of CBR in the phases of retrieval (section 4.3.1.1) and reuse (section 4.3.1.2) for testing purposes. Figure 5.5 presents an example of the matching procedure when the initial knowledge domain model has been established. As can be seen, the matching of findings during the retrieval phase are listed as direct, partial or unmatched. A direct match occurs when an arc to the same finding (concept) is present in both the unsolved case and the retrieved solved case. The partial match indicates that a match can be found using inference on the parts of the semantic network where both cases are integrated. If there is a way to traverse to the same finding through other concepts (for instance by using inheritance), the match is therefore described as partial. If no match can be found, it is listed as unmatched.

The retrieval phase presents the most similar cases to the unsolved case, ranked by the total amount of direct and partial matches. When entering the reuse phase, an initial attempt is to transfer the solution of the solved case to the unsolved case.

The analysis within a domain knowledge model is heavily based on methods of inference. Inference in jCreek is based on two methods, the activation spreading and inheritance. The activation spreading is initiated at one concept and the activation is spread recursively along its arcs to its neighbours. A spreading activation control mechanism specifies which relations to traverse and the maximum search depth. Inference using inheritance is a special case of the activation spreading, where concepts are allowed to inherit relationships from other concepts along a restricted set of arcs ('subclass-of', 'instance-of'). With the traditional object oriented inheritance of the Java language, it is possible to extend these methods further within jCreek.

The current initial model is based on the use of single standard view planes during analysis. Complex abnormalities might require further investigations of its presence in several view planes. Reasoning based on several standard view planes at the same time is described in chapter 7.

5.3 Workflow

This section will describe the workflow between the components elaborated in the previous sections. The input and output of each component are explained. The complete workflow is illustrated in figures 5.6 and 5.7, which defines the structure of this section. Elliptical shapes with solid lines represent components, while the elliptical shapes with dashed lines are subcomponents. Subcomponents are procedures needed to successfully combine the components, as will become apparent through the discussion in this section. The retrieve, reuse, revise, and retain solution components shown in figure 5.7 indicate the four phases of the CBR cycle, as discussed in section 4.3.1.

The initial input is the standard view plane extracted during the examination, represented as a matrix of image pixels. The region extraction component takes the matrix of image pixels as input and aims to extract the regions of interest. The regions present in each image are segmented and the output is a matrix of image pixels with the same dimensions as the one

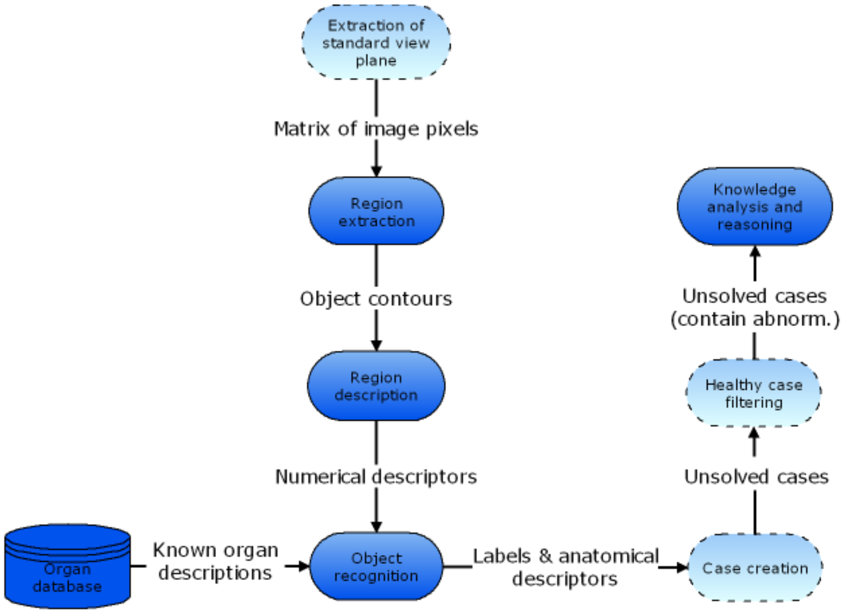


Figure 5.6: Illustration of the first three components, and the subcomponents needed to begin knowledge analysis and reasoning.

given as input, but with the additional labeling of the contours used for later processing.

After extracting the region contours from the image, the next component performs region description. The contours, and hence the regions, are described numerically, enabling further comparison. The output of this component is a vector of numerical descriptors representing each region in the original image.

The third component of the pipeline handles object recognition. Numerical descriptors created by the previous component are matched against descriptions of known organs and symptoms. During this procedure, each region is given a particular label, indicating the most similar organ already described in what is referred to as the organ database (see section 5.2.3). In addition to the labels assigned to each region, the anatomical descriptors discussed in chapter 2, are obtained during this procedure. The total output equals the labels and the anatomical descriptors.

The previous output is transformed into an unsolved case applicable to a CBR system within the case creation subcomponent. The transformation is performed by converting the knowledge obtained in the previous step into the representation language used by the CBR system, using the mapping mentioned in 5.2.4.

Every unsolved case is processed through the healthy case filtering mentioned in section 5.2.4. This subcomponent aims to reduce the amount of computation applied to cases which only confirm a healthy fetus. The input is the complete collection of unsolved cases, while the output is a filtered set of cases containing symptoms of abnormalities.

After filtering the unsolved cases, the cases are being analysed. This component is further decomposed in figure 5.7. Here, the general domain knowledge is obtained and created through the process of domain knowledge organisation. When presented with an unsolved case, a similar case is retrieved from the case base in the retrieval phase of the CBR cycle. Both the unsolved case and the newly retrieved case are given as input to the reuse phase. Based on the unsolved case, and the retrieved case, they are combined to form an initial solution to the problem presented in the unsolved case. This solution is suggested to the user while the newly created solved case is sent to the revise phase for further evaluation. This phase is responsible for testing the validity of the solution. If the validity check fails, the solution must be corrected accordingly. This phase often depends on user interaction, which might create a bottleneck in the CBR system. In the last phase, the repaired case is retained and integrated into the case base. During the integration, the case can be added as a new entry in the case base, or used to generalise an existing case to incorporate the newly discovered knowledge. By storing the knowledge in the database, it can be utilised during the next analysis of a similar standard view plane.

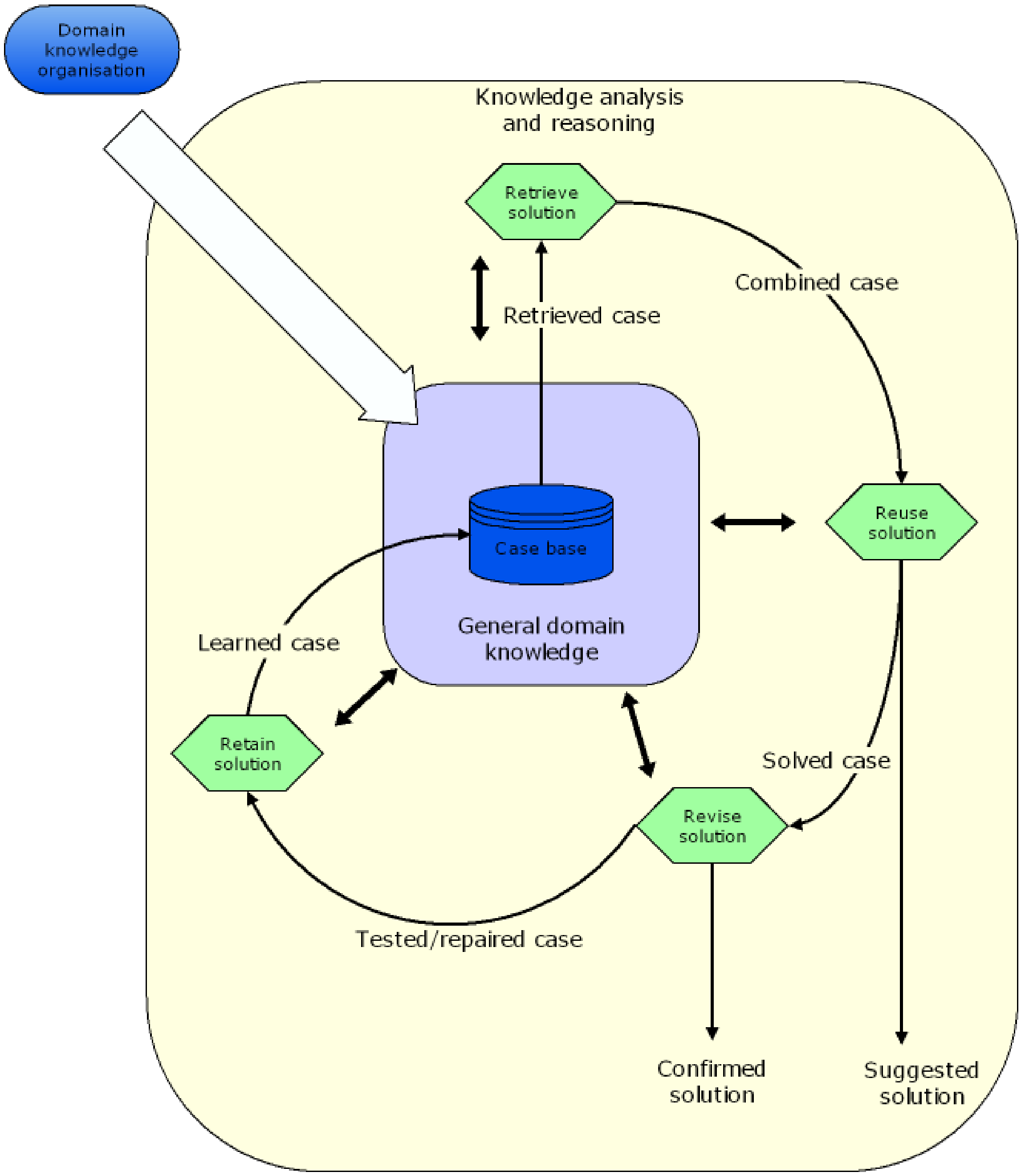


Figure 5.7: The decomposition of knowledge analysis and reasoning, and its interaction with domain knowledge organisation.

Chapter 6
Evaluation

This chapter will present an evaluation of the work performed in this thesis, based on the main task and subtasks described in chapter 1. The evaluation of each subtask presents contributions made by this thesis and challenges yet to be solved before the prototype can be applied in a routine ultrasonography.

6.1 Standard view planes and features

A foundation for using standard view planes during the routine ultrasonography has been presented in [1]. This thesis has further explained and refined a subset of these standard view planes. The reason for using only a subset is to ensure that the description of each standard view plane is as accurate as possible. Because of insufficient training in the medical domain, the creation of a complete description of every standard view plane is beyond the scope of this thesis. The refinement of the standard view planes presented in [1] should be performed in close cooperation with medical domain experts.

The description of each standard view plane is based on the anatomical landmarks (representing the normalities present), anatomical descriptors (describing quantitative measurements), and symptoms of abnormalities present in the particular standard view plane. This ensures that both healthy cases and cases containing abnormalities are found. Due to the complexity of eliciting medical knowledge, only a subset of the symptoms indicating a certain abnormality is presented.

To enable real world application of standard view planes, a thorough elaboration beyond the work presented in this thesis is required. The number of standard view planes should be sufficiently large to cover most of the fetal body and the anatomical landmarks, anatomical descriptors, and symptoms present in every standard view plane should be further extended.

6.2 Feature extraction and understanding

This section will evaluate the task of extracting and understanding features from an ultrasound image. Although the task seems to consist of two phases: extraction, and matching, they are closely interconnected, implying a common evaluation.

The first step towards feature extraction is the segmentation phase, which was performed using the live-wire approach. It offers reliable results with medical expert guidance, and it does not suffer from the severe time limitations imposed by a completely manual segmentation routine. The main drawback of this method is its ineptness for full scale implementation due to manual intervention. If the standard view planes should be analysed in real time during the routine examination, an automatic segmentation routine must be implemented. A promising example of such a procedure was presented in section 3.3.1.4.

The description created for each segmented region is based on the region descriptors presented in section 3.5.3. The descriptors fulfill the requirements related to invariance of ultrasound images (see section 3.5.3). If these requirements were not fulfilled, the subsequent matching procedure would have to be altered.

The matching procedure, as presented in this thesis, is based on the creation of an organ database. Each record in the database represents an already described organ in a particular standard view plane. This ensures an easy way of extracting this knowledge whenever a new organ requires labeling. By structuring the database according to standard view planes, only organs belonging to the same standard view plane are considered when a new region is found. As stated in section 5.2.3.1, there should be no need to incorporate for instance a CBR system when recognising regions. This eases the processing of the image, and reduces the time spent before a diagnosis can be made.

A negative aspect of the organ database, is its lack of a standardised procedure for adding new information. This is currently performed manually after consulting a medical expert, implying an almost static database. In a complete system, the maintenance of the database should be designed to offer a more dynamic approach. In addition, the database must also be extended compared to its current contents. The focus has been to create a working prototype, which by no means contains enough information to recognise every region in the standard view planes. This is a comprehensive task, which requires a lot of cooperation with medical experts to ensure the correctness of the data.

An important issue to consider is the degree of connectivity between the parts constituting the task of feature extraction. The design proposed in this thesis consider each component as a single unit creating an output for the next unit. The advantage of such a procedure is that each component can be easily replaced without altering its surroundings. The negative aspect of the current approach, is that a series of steps is necessary before arriving at the conclusion. A better approach would be to have a surrounding framework controlling the different components.

6.3 Feature analysis

The feature analysis of this thesis has been based on the Creek architecture for case-based reasoning. An initial domain knowledge model containing both general domain knowledge and knowledge resulting from the understanding of ultrasound images has been created. The domain knowledge model was created using the semantic networks of Creek and two dimensions were initially added. As a proof-of-concept, the model indicates the range of complexity concerning the medical domain, but cannot account for complex relations required by the domain. As such, the model should be further extended and refined before being used in a medical application.

Most ultrasound images analysed are confirmations of a healthy fetus, and should not require excessive amounts of processing. An approach of filtering healthy cases has been proposed by this thesis, but other techniques should also be considered.

Before implementing the procedures for analysis, it is important to consider the need of a user interface. A complex reasoning and learning will not be possible without the feedback and confirmation from the user. This thesis has made no attempt to suggest the manner in which this should be done, but recommends a thorough examination of this topic. A few guidelines are given in section 7.1.1.

The main focus of this thesis has been the retrieval and early reuse stage of the CBR-cycle. Incorporation of the complete cycle would require work on issues such as the integration of learned cases, and an effective interface for user influence on solution development.

Chapter 7
Future work

The work performed in this thesis is an initial attempt at solving the problems presented in chapter 1. The designed components and the implemented prototype do not represent a complete solution to these problems, but rather a work in progress to analyse the challenges presented and the complexity of the medical domain. A lot of research remains and this chapter will try to point out some important areas of future work, apart from the further enhancement of those already described in this thesis, especially the expansion of the medical domain knowledge model.

Several important issues have not been included during the design and development of the prototype described in this thesis. The prototype will not operate in an isolated environment and its interaction with other systems as well as users must be considered. Section 7.1 will elaborate further on this matter. The second issue addressed is an extension of the discussion of the analysis of single versus multiple standard view planes when making a diagnosis, as presented in chapter 5. A majority of the abnormalities have symptoms present in several standard view planes and this should be accounted for. A possible approach to mend this is described in section 7.2.

7.1 Interaction

This thesis has made no attempt to explain the communication between the initial prototype and its natural environment of users and ultrasound equipment. To enable the actual use of a prototype such as the one suggested in this thesis, a lot of attention must be given to the issues of interaction. The two main issues are interaction with the user (midwife), and the integration of a prototype within an environment of modern ultrasound equipment. Section 7.1.1 will focus on the interaction between the user and the prototype, while section 7.1.2 will consider the integration of a prototype within a working ultrasound environment.

7.1.1 User interaction

To be able to utilise the research of this thesis, the suggested prototype must be able to interact with the user performing the routine ultrasonography, namely the midwife. As mentioned in chapter 1, the work presented has focused on aiding less experienced midwives. This imposes a number of possible requirements to be considered:

- **Feedback**
  Considering the fact that the midwife performing the routine ultrasonography has less

experience using the ultrasound equipment, feedback is very important. Feedback can be manifested as a small indication of the current image status, an enquiry of a more detailed examination based on earlier experiences, or the need to consult a medical expert. The feedback should be presented to the midwife within the working user interface in a discrete way to avoid unnecessary distraction.

- **Seamless integration**
  To avoid confusion and unneeded distraction, the prototype should be seamlessly integrated with the systems currently known to the midwife. No extra complexity should be induced, except from the introduction of standard view planes described in chapter 2. The focus of the education should not change, but rather take advantage of the remedies provided by the introduction of the prototype functionality.

- **Logging**
  With the prototype analysing every standard view plane of a routine ultrasonography, a procedure for logging the results and discoveries would be very helpful, if not absolutely necessary. To avoid distraction, only a small subset of the output from the prototype should be presented to the user. With the introduction of extensive logging, the medical record of the fetus being examined can be significantly extended. This grants the possibility of further examination of features that was not discovered during the examination.

7.1.2 System integration

In addition to the interaction with a user, the modules of the prototype must be integrated within a current working environment of ultrasound software and hardware. To become a useful addition to the existing equipment, a number of technical issues such as speed, flexibility and extensibility have to be addressed:

- **Speed**
  To increase the usability of ultrasound equipment containing the modules of the prototype designed in this thesis, speed is an important matter. In this context, speed is essentially a measurement of how fast the modules are executed during normal use. The integration of new modules should not yield an overall slower system, but rather be a useful addition without creating visible limitations to the user. Practically, the decision of whether to integrate the prototype into hardware or software within the existing equipment, might have severe impacts on the execution speed.

- **Flexibility**
  Flexibility denotes the level at which the prototype modules are integrated with the modules of the existing ultrasound system. Two main possibilities should considered; integrating the modules as internal, or external components. Internal integration requires a complete specification of the communication with the relevant modules of the existing system. Using external integration of modules, a general interface for connecting to the other parts of the system will suffice. The most significant drawback of an external integration is the loss of execution speed caused by the bottleneck at the communication interface, while one of its benefits is the ease of replacement.

- **Extensibility**
  An important property of a prototype such as the one described in this thesis, is extensibility. As the medical domain is an extensive and complex domain, the ability to further expand models and procedures of the designed modules becomes very important. The extensibility of the prototype modules depends on the choices made on the matters of speed and flexibility. The extension of an internal hardware module requires quite a different approach than the extension of an external software module. These are all issues requiring consideration before integrating the prototype with the existing software and hardware of current ultrasound equipment.

7.2 Session-based cases

Analysis and reasoning based on several standard view planes instead of a single one would surely increase the accuracy of a diagnosis. A possible approach to include several view planes into a single retrieval phase, is to introduce a session case. The session case contains several of the most important view planes acquired during a routine ultrasonography. It includes significantly more knowledge than the single cases discussed in chapter 5. An illustration of the session case is shown in figure 7.1. The figure shows the similarity assessment of an unsolved session case with a solved case. The session case is composed of several instances extracted from the image processing (understanding) procedure. Each instance resembles the extraction of image knowledge from a single standard view plane[1]. The main idea is that a combined set of findings yields a match of greater accuracy than the match of a single instance.

The session case resembles the work of a human expert in a better way than the single case processing. A human expert would consult the knowledge from different parts of the fetus to ensure a satisfactory reliability of the diagnosis. The reason and analysis approaches presented in chapter 5 will still apply to the session case, without changing the underlying framework. A procedure for removing duplicate knowledge from the session case should be designed to avoid unnecessary overhead during the similarity assessment in the retrieval phase.

[1]Unsolved case #3 is not the findings of a standard view plane, but additional information that may affect the diagnosis.

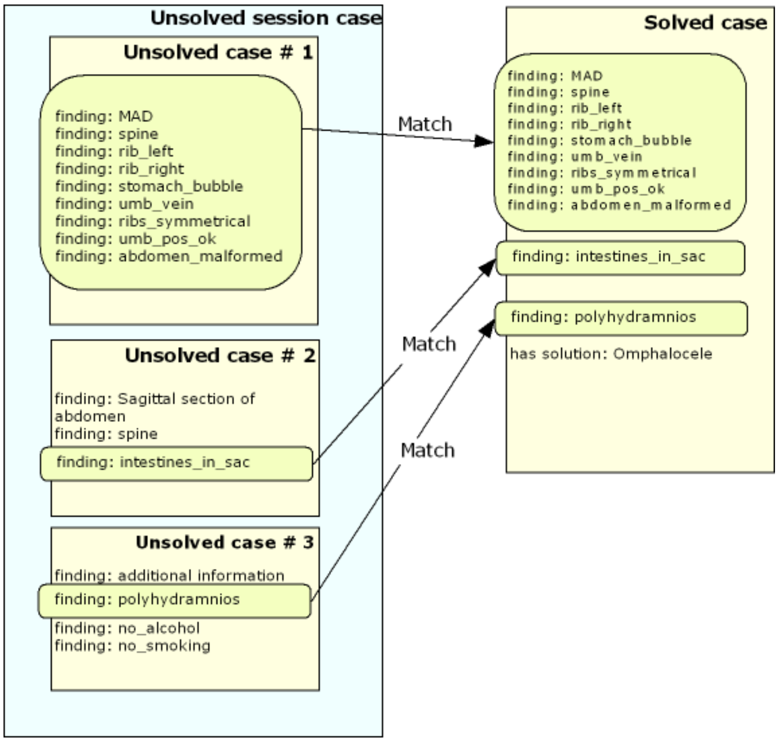


Figure 7.1: Case retrieval using session cases.

Appendix A
Live-wire

As mentioned in section 3.7, the live-wire approach is chosen as segmentation procedure for this thesis. This appendix will explain this method more in detail [1]. The principles of the procedure will be shortly restated, before explaining the more specific details of the procedure. These are in particular the generation of the cost matrix (see section A.1) and the generation of the optimal path (see section A.2).

The live-wire approach is a semi-automatic segmentation procedure using graph search to estimate the contour. Each pixel in the image is treated as a node in the graph. The arcs of the graph are the costs calculated for each pixel and its neighbours. The procedure starts by assigning control points along the contour of a region of interest. The contour is then created by finding the shortest path between adjacent control points.

A.1 Cost matrix

The live-wire approach is dependent on the generation of a cost matrix in order to obtain the optimal path between two points. The basis for this cost matrix is the image gradient. In order to incorporate both the direction and the magnitude of the gradient, Mortensen et al. [31] combine these two measurements into a single matrix using the following equations:

$$G(x,y) = \sqrt{{G_x}^2(x,y) + {G_y}^2(x,y)}$$

$$O(x,y) = tan^{-1}\left(\frac{G_x(x,y)}{G_y(x,y)}\right)$$

$$c(x,y) = [max_{x,y}(G(x,y)) - G(x,y)] + \alpha \mid O(x,y) * f - O(x,y) \mid$$

where G_x and G_y represents the horizontal and vertical gradients of the image. The first term of $c(x,y)$ is minimized whenever the gradient magnitude, $G(x,y)$, has a high value. The second term includes a smoothing term, f, which can be either an averaging filter or a Gaussian filter, and a scaling factor, α. By computing the absolute difference of the smoothed orientation image, $O(x,y)$, and itself, the cost is minimised whenever the direction of the gradient orientation of one pixel is similar to the orientation of the surrounding gradients. A single matrix, $c(x,y)$, has now been obtained, with low edge cost corresponding to a strong gradient magnitude. Gradient orientation similar to its neighbours will further decrease the

[1]The information in this chapter is largely obtained using the results given in [31].

cost along the specified path. To compensate for the difference in horizontal and diagonal distance, a weighting scheme is also associated to the cost matrix. Here, diagonal connections have a weighting which is $\sqrt{2}$ times larger than horizontal and vertical distances. This ensures that the correct Euclidean distance is maintained. If there were no weights assigned, the cost along a diagonal path would be too low compared to the correct distance.

A.2 Optimal path

Once the cost matrix has been constructed, the contour can be created by finding the optimal path between the control points. In the original paper, the path is not computed between specific points, but rather from one starting point to all other points in the graph. The approach used is Dijkstra's algorithm, first presented in [52]. The algorithm starts at a specified starting point, and estimates the cost to all its neighbours. Then the neighbour with the lowest cost is chosen, and the cost to the neighbours of this node is estimated. This process is repeated until the distance from the starting point to all other nodes in the graph has been estimated. In addition to the cost of reaching the node, a pointer is stored in each node pointing to the previous node in the path. By storing the pointers, a path can be found from each point leading back to the original starting point. This path will then, because of the cost matrix, follow edges in the image. Figure A.1 illustrates the concepts of both the cost matrix and the optimal path algorithm. Figure (a) shows the initial cost matrix. The low costs indicates an edge in the matrix. As can be seen when expanding the graph in figures (b)-(d), the cost to reach each pixel from the starting pixel is calculated. The lines between each pixel indicate the path to the pixel. In figure (e), all pixels have been expanded, and the matrix shows the cost of reaching each pixel in the image and also the path leading from each pixel back to the original pixel.

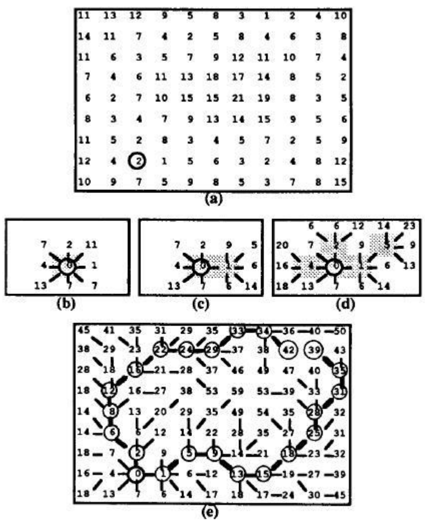


Figure A.1: Illustration of cost matrix and generation of optimal path. Adapted from [31].

Appendix B
Descriptors

As mentioned in section 5.2.2, several descriptors have been used to numerically describe the shapes of the objects in the ultrasound image. This chapter explains the descriptors of choice in detail, and provide their mathematical formulation.

B.1 Moments

An image is often considered as a two-dimensional function, where the input of the two coordinates gives the brightness value of the image at a particular point. By normalising this function, it can be seen as a probability density of a two-dimensional random variable. The moments are defined to be the function describing the properties of this random variable [36]. These moments are used to describe regions of objects on a black background. The original moment of order (p+q) is defined as:

$$m_{pq} = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} i^p j^q f(i,j) \tag{B.1}$$

where i and j represent the pixel coordinates of the image. This moment, often referred to as raw moment, suffers from its lack of invariance. It is dependent on both translation, rotation, and scaling. It must be modified in order to obtain invariance to all three transformations. Translation invariance is achieved by using the centroid of the region as a starting point. The centroids are defined as:

$$\begin{aligned} x_c &= \frac{m_{10}}{m_{00}} \\ y_c &= \frac{m_{01}}{m_{00}} \end{aligned} \tag{B.2}$$

which yields the central moment of order (p+q):

$$\mu_{pq} = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (i - x_c)^p (j - y_c)^q f(i,j) \tag{B.3}$$

After obtaining the central moment, the scale-invariant moment of order (p+q) is obtained by the following equation:

$$\eta_{pq} = \frac{\mu'_{pq}}{(\mu'_{00})^{\gamma}} \tag{B.4}$$

where

$$\gamma = \frac{p+q}{2} + 1 \tag{B.5}$$

and

$$\mu'_{pq} = \frac{\mu_{pq}}{\alpha^{(p+q+2)}} \tag{B.6}$$

Before obtaining moments invariant to all three transformations, the unscaled central moments of order (p+q), ϑ_{pq}, must be obtained:

$$\vartheta_{pq} = \frac{\mu_{pq}}{(\mu_{00})^{\gamma}} \tag{B.7}$$

Finally, rotation-invariant moments can be obtained if the coordinate system is chosen such that $\mu_{11} = 0$. There are a total of seven invariant moments, which are defined as:

$$\varphi_1 = \vartheta_{20} + \vartheta_{02} \tag{B.8}$$

$$\varphi_2 = (\vartheta_{20} - \vartheta_{02})^2 + 4\vartheta_{11}^2 \tag{B.9}$$

$$\varphi_3 = (\vartheta_{30} - 3\vartheta_{12})^2 + (3\vartheta_{21} - \vartheta_{03})^2 \tag{B.10}$$

$$\varphi_4 = (\vartheta_{30} + \vartheta_{12})^2 + (\vartheta_{21} + \vartheta_{03})^2 \tag{B.11}$$

$$\begin{aligned} \varphi_5 = {} & (\vartheta_{30} - 3\vartheta_{12})(\vartheta_{30} + \vartheta_{12})[(\vartheta_{30} + \vartheta_{12})^2 - 3(\vartheta_{21} + \vartheta_{03})^2] \\ & + (3\vartheta_{21} - \vartheta_{03})(\vartheta_{21} + \vartheta_{03})[3(\vartheta_{30} + \vartheta_{12})^2 - (\vartheta_{21} + \vartheta_{03})^2] \end{aligned} \tag{B.12}$$

$$\varphi_6 = (\vartheta_{20} - \vartheta_{02})[(\vartheta_{30} + \vartheta_{12})^2 - (\vartheta_{21} + \vartheta_{03})^2] + 4\vartheta_{11}(\vartheta_{30} + \vartheta_{12})(\vartheta_{21} + \vartheta_{03}) \tag{B.13}$$

$$\begin{aligned} \varphi_7 = {} & (3\vartheta_{21} - \vartheta_{03})(\vartheta_{30} + \vartheta_{12})[(\vartheta_{30} + \vartheta_{12})^2 - 3(\vartheta_{21} + \vartheta_{03})^2] \\ & - (\vartheta_{30} - 3\vartheta_{12})(\vartheta_{21} + \vartheta_{03})[3(\vartheta_{30} + \vartheta_{12})^2 - (\vartheta_{21} + \vartheta_{03})^2] \end{aligned} \tag{B.14}$$

B.2 Area and circumference

Area and circumference are two of the simplest ways of describing a region, but as mentioned in section 5.2.2, they are both invariant to rotation and translation, making them a suitable choice when describing regions in a medical context. Area is defined to be the number of pixels constituting the region. The descriptor is made scale-invariant by computing the percentage of the total pixels contained in each region. The output of this descriptor is thus a percentage value between zero and one.

Circumference is defined as the number of pixels along the border of the region. When computing the region, and hence also the circumference, 8-connectedness between the pixels is used. The circumference can also be adjusted to be scale-invariant, just as the area descriptor.

B.3 Compactness

This descriptor is used to describe the degree of compactness in a particular region. It combines the area and the circumference and is defined to be:

$$compactness = \frac{circumference^2}{area} \tag{B.15}$$

A region is said to have a high degree of compactness whenever this descriptor has a low value. The most compact region is the circle, while more irregular shapes will have a higher value and hence are said to be less compact. Since the circumference is a combination of area and circumference, its degree of invariance is the same as that of area and circumference.

Bibliography

[1] Jan Ødegård and Anders Østen. *Knowledge based medical image analysis*, 2005.

[2] Sturla H. Eik-Nes. *Vekstkontroll-målinger*.

[3] Norman C. Smith. *Obstetric ultrasound made easy*. Harcourt Publishers Limited, 2002. ISBN: 0443-07013X.

[4] Joseph Woo. *Obstetric Ultrasound*. http://www.ob-ultrasound.net/.

[5] Dominic A. Carone. *MedFriendly.com*. http://www.medfriendly.com/lateralventricles.html, 2006.

[6] Alexandra Novakov. *Ultrasound examination in the second and third trimester*. http://www.thefetus.net/page.php?id=449, 2002.

[7] Gianluigi Pilu and Kypros Nicolaides and Renato Ximenes and Philippe Jeanty. *Diploma in Fetal Medicine and ISUOG Educational Series*. http://www.centrus.com.br/DiplomaFMF/SeriesFMF/18-23-weeks/index-18.htm, 2000.

[8] Roger C. Sanders. *Structural Fetal Abnormalities*. Mosby, 2002. ISBN: 0-323-01476-3.

[9] Douglas S. Richards and Gene Cornwall. *Abdominal Circumference*. http://www.obgyn.ufl.edu/ultrasound/MedinfoVersion/sec6/6_4.html, 1998.

[10] Joseph S.K. Woo. *The Abdominal and Thoracic circumference – some historical notes*. http://www.ob-ultrasound.net/ac_historical.html, 1998.

[11] R.B. Kalish, S.T. Chasen, L. Rosenzweig, and F.A. Chervenak. Esophageal atresia and tracheoesophageal fistula: the impact of prenatal suspicion on neonatal outcome in a tertiary care center. *Journal of Perinatal Medicine*, 31:111–114, 2003.

[12] Rovida Roberto. *TRV Fetal Abdomen, Duodenal Atresia (Double-Bubble sign)*. http://www.obgyn.net/us/us.asp?page=gallery/gallery.

[13] E. Tegnander, S. H. Eik-Nes, and D. T. Linker. Incorporating the four-chamber view of the fetal heart into the second-trimester routine fetal examination. *Ultrasound Obstet. Gynecol.*, 4:24–28, 1998.

[14] Sturla H. Eik-Nes, Harm-Gerd Blaas, and Eva Tegnander. *Nicer yearbook*, 1996.

[15] David A. Nyberg, Barry S. Mahony, and Dolores H. Pretorius. *Diagnostic Ultrasound of Fetal Anomalies*. Mosby - Year Book Inc., 1990. ISBN: 0-8151-6439-4.

[16] A.D.A.M. Inc. *A.D.A.M. Inc :: Complete Online Health Content Solutions*. http://www.adam.com/, 2005.

[17] Bjørn A. J. Angelsen. *Ultrasound Imaging - Waves, signals and signal processing, Vol I*. Emantec AS, 2000. ISBN: 82-995811-0-9.

[18] Atlantic Interactive. *Basic principles of ultrasound*. http://www.echo-web.com/html/echo-202-free/echo202-1-body.asp?code=, 2004.

[19] Øystein Svinning. Segmentation of objects in 2D and 3D ultrasound images. Master's thesis, Norwegian University of Science and Technology, 2003.

[20] Michael Kass, Andrew Witkin, and Demetri Terzopoulos. Snakes: Active Contour Models. *International Journal of Computer Vision*, pages 321–331, 1988.

[21] Lilian Ji and Hong Yan. Attractable snakes based on the greedy algorithm for contour extraction. *Pattern Recognition*, 35:791–806, 2002.

[22] Chenyang Xu and Jerry L. Prince. Snakes, Shapes, and Gradient Vector Flow. *IEEE Transatctions on Image Processing*, 7(3):359–369, 1998.

[23] D.J. Williams and M. Shah. A Fast Algorithm For Active Contours And Curvature Estimation. *Image Understanding*, 55:14–26, 1992.

[24] P. Abolmaesumia and M.R. Sirouspourb. Ultrasound Image Segmentation Using an Interacting Multiple Model Probabilistic Data Association Filter. *Proceedings of SPIE International Conference on Medical Imaging*, 5370:484–493, 2004.

[25] Chung-Ming Chen, Henry Horng-Shing Lu, and Yu-Chen Lin. An Early Vision-Based Snake Model for Ultrasound Image Segmentation. *Ultrasound in Medicine and Biology*, 26(2):273–285, 2000.

[26] Chung-Ming Chen, Henry Horng-Shing Lu, and An-Ting Hsiao. A Dual-Snake Model of High Penetrability for Ultrasound Image Boundary Extraction. *Ultrasound in Medicine and Biology*, 27(12):1651–1665, 2001.

[27] G. Hamarneh and T. Gustavsson. Combining Snakes and Active Shape Models for Segmenting the Human Left Ventricle in Echocardiographic Images. *IEEE Computers in Cardiology*, 27:115–118, 2000.

[28] Jon Shlens. *A tutorial on principal component analysis*, 2003.

[29] Judith G. Thomas, Richard Alan Peters II, and Philippe Jeanty. Automatic Segmentation of Ultrasound Images Using Morphological Operators. *IEEE Transactions on Medical Imaging*, 10(2):180–186, 1991.

[30] Greg Welch and Gary Bishop. *An Introduction to the Kalman Filter*, 2004.

[31] E. Mortensen, B. Morse, W. Barrett, and J. Udupa. Adaptive boundary detection using live-wire two-dimensional dynamic programming. *Proceedings of Computer in Cardiology*, pages 635–638, 1992.

[32] H. Ji, A. Shen, G. Wang, and X. Gao. An Interactive Segmentation Method for Medical Images. *6th International Conference on Signal Processing*, 1:580–582, 2002.

[33] Jiankang Wang and Xiaobo Li. A System for Segmenting Ultrasound Images. *Proceedings of the Fourteenth International Conference on Pattern Recognition*, 1(1):456–461, 1998.

[34] N. Hiransakolwong, K. A. Hua, K. Vu, and P. S. Windyga. Segmentation of Ultrasound Liver Images: An Automatic Approach. *Proceedings of the ICME International Conference on Multimedia and Expo*, 1(1):573–576, 2003.

[35] Rafaen C. Gonzalez and Richard E. Woods. *Digital Image Processing*. Prentice Hall, second edition, 2002.

[36] M. Sonka, V. Hlavac, and R. Boyle. *Image processing, analysis, and machine vision*. PWS Publishing, second edition, 1999.

[37] Agnar Aamodt. Knowledge-Intensive Case-Based Reasoning and Sustained Learning. *ECAI-90,Ninth European Conference on Artificial Intelligence*, pages 1–6, 1990.

[38] Agnar Aamodt. *A knowledge-intensive, integrated approach to problem solving and sustained learning*. PhD thesis, University of Trondheim, Norwegian Institute of Technology, Department of Computer Science, 1991.

[39] Agnar Aamodt and Enric Plaza. Case-based Reasoning; Foundational Issues, Methodological Variations, and System Approaches. *AI Communications*, 7(1):39–59, 1994.

[40] Agnar Aamodt. Explanation-Driven Case-Based Reasoning. *In S. Wess, K. Althoff, M. Richter (eds.): Topicsin Case-based reasoning*, pages 274–288, 1994.

[41] Agnar Aamodt. Knowledge-intensive case-based reasoning in Creek. *Advances in case-based reasoning, 7th European Conference, ECCBR 2004, Proceedings*, 7(1):39–59, 2004.

[42] R. Schmidt, S. Montani, R. Bellazzi, L. Portinale, and L. Gierl. Case-Based Reasoning for medical knowledge-based systems. *International Journal of Medical Informatics*, 64(2-3):355–367, 2001.

[43] Ketil Bø and Agnar Aamodt. A new approach to applicable memory-based reasoning. *Journal of Experimental and Theoretical AI*, 11:479–496, 1999.

[44] Bruce W. Porter, Ray Bareiss, and Robert C. Holte. Concept learning and heuristic classification in weak-theory domains. *Artificial Intelligence*, 45(1-2):229–263, 1990.

[45] Agnar Aamodt. *Lecture notes IT2702 Artificial Intelligence*, 2005.

[46] M. Egmont-Petersen, D. de Ridder, and H.Handels. Image processing with neural networks - a review. *Pattern Recognition*, 35(10):2279–2301, 2002.

[47] Bart Kosko and Satoru Isaka. *Fuzzy Logic*. `http://www.fortunecity.com/emachines/e11/86/fuzzylog.html`, 1993.

[48] Alexandre R.J. François and Gérard Medioni. Generic Shape Learning and Recognition. *Proceedings of the International Workshop on Object Representation in Computer Vision II*, pages 287–320, 1996.

[49] Kristian Bendiksen. Case-Based Image Interpretation. Master's thesis, Norwegian University of Science and Technology, 2004.

[50] A. Th. Schreiber and B. J. Wielinga. *Knowledge Model Construction*. `http://ksi.cpsc.ucalgary.ca/KAW/KAW98/schreiber/`, 1998.

[51] Department of Computer and Information Science. *Creek - Case Based Reasoning*. `http://creek.idi.ntnu.no/`, 2005.

[52] Edsger W. Dijkstra. A note on two problems in connexion with graphs. *Numerische Mathematik*, 1:269–271, 1959.