NTNU

Norwegian University of
Science and Technology

Optical flow applied to infant movement

Harald Kirkerød

Master of Science in Engineering Cybernetics
Submission date: January 2010
Supervisor: Øyvind Stavdahl, ITK
Co-supervisor: Annette Stahl, ITK

Norwegian University of Science and Technology
Department of Engineering Cybernetics

Problem Description

According to the GMA method, characteristic movements in this period are highly correlated to healthy motor development, while absence of these movements is a strong indication of cerebral pathology that may result in cerebral palsy.

An existing database contians video footage of approx. 100 infants, recorded a few weeks after birth. In this assignment you will apply modern methods based on optical flow and dynamic modeling techniques to detect movement pattern differences between infants with normal motor development and those with an abnormal development.

1. Review the literature regarding optical flow and related methods. Give examples of previous use of these methods for human motion analysis, and point out possible drawbacks or advantages of these solutions.

2. Research and describe at least one complete analysis set-up that includes optical flow based preprocessing and subsequent classification.

3. Design and implement a suitable software system that incorporates the adapted method(s) and can be used for assessing the method(s) based on the available database.

4. To the extent possible within the given time limits, perform an actual data analysis and assess the system's performance in light of previous results.

Preferably, the report should be structured as a draft scientific report and an additional part providing introductory information and other relevant material.

Assignment given: 24. August 2009
Supervisor: Øyvind Stavdahl, ITK

Acknowledgements

Personally I found the project rewarding in both its cause and with its openness for choosing my own research paths. It has also strenghtened my own self-confidence as a technologist. I would like to thank my supervisor Øyvind Stavdahl for his wholehearted interest in the project and his ability of always boosting my motivation in our meetings. My gratitude also goes to my co-advisor Annette Stahl, for her well of expertise and ideas, and many pleasent talks.

Finally I would like to praise to the skies my dear Ana for invaluable support throughout the project.

Harald Kirkerød
January 24, 2010

Abstract

Healthy infants in the age group of 9 to 20 weeks post term have a distinct movement pattern called fidgety movements. The abscense of, or anomalies in these movements are indications that the infant may suffer from Cerebral Palsy. Data from video clips of infants have been extracted and used as features for classifying these movements as normal or abnormal. An earlier project used the concept of motiongrams for extracting the centroid of motion and quantity of motion of the image from each frame of the video clips. This project attempts to improve the dataset by extracting the same features using the concept of optical flow.

Using the classification methods that yielded the best results from the motiongram project, the new dataset based on optical flow was constructed and tested for comparison. The image was as before divided into four quadrants with a circle of varying radius in the middle. Tracking how the centroid of motion moved between the separated areas of the image generated a transition probability matrix. By classifying on the entropy and variance of this matrix, a sensitivity of 90,0% and a specificity of 94,6% was reached, which was an improvement compared to the best classification result based on motiongrams with a sensitivity of 90,0% and specificity of 86,6%. The overall performance of both methods are around a sensitivity of 90% and a specificity of 70%.

Using motiongrams and optical flow to extract features from video data for classification has turned out to be a promising approach for a simple and non-invasive objective tool for diagnosing infants with Celebral Palsy. The movement information in an optical flow field does however have a much higher potential for generating discriminative features than just as the simple representation a centroid of motion actually is.

Contents

Nomenclature

CoM Centroid of Motion

CP Celebral Palsy

GMA General Movement Assessment

LD Linear discriminant function

LD Quadratic discriminant function

LDC Linear discriminant function with diagonal covariance matrix estimate

MDD Mahalanobis distances discriminant function with stratified covariance estimates

MM Markov model transition matrix

NaN Not a Number

NM Normalized transition matrix

QDC Quadratic discriminant function with diagonal covariance matrix estimate

QoM Quantity of motion

QoMC Quantity of Motion Change

SE Sensitivity

SP Specificity

1 Introduction

Today a large amount of training is needed for a clinician to learn the General Movement Assessment method, which is a subjective observation method. Thus a more objective assessment solution could be beneficial.

An earlier project [18] extracted features from an existing set of video data of infants with and without the diagnosis Celebral Palsy using a rather simple approach. These features were then classified with promising results, which indicated that the method could be of assistance with diagnosing the infants.

The purpose of this project is to use a more advanced method based on optical flow for extracting the same features of the video set as for the earlier project using motiongrams. A comparison of the two methods will then be assessed together with other classification methods based on the new dataset.

Chapter 2 contains background material for the diagnosis Cerebral Palsy and the General Movement Assessment method. A short introduction to research attempts on infant movement assessment, optical flow for human motion analysis and other relevant work will also be given.

In Chapter 3, a short description of the video data set is provided, followed by the theory of motiongrams which was used in the earlier project. Then the theory behind optical flow is explained, along with a short explanation of Markov models.

Chapter 4 describes the methods for this thesis, starting with the preprocessing of the video data. An explanation of how the dataset is aquired is given, followed by the feature extraction process and the filtering of the dataset. The chapter ends with a short explanation of how the classification was performed.

The results of the comparison between the earlier project and this thesis are given in Chapter 5, followed by a discussion of the results in Chapter 6.

2 Background

This chapter provides the background on the subjects in matter for this project. The related work is divided into the three sections. The first one deals with other projects related to infant movement assessment. Optical flow used in human motion analysis is in the second part, while other related work dealing with human motion analysis is found in the last part. The chapter ends with a short summary of the related work.

2.1 Cerebral Palsy (CP)

Cerebral Palsy (CP) is defined as [29]:

> Cerebral palsy is a general term referring to defects in motor function or coordination resulting from several types of brain damage, which may be caused by abnormal brain development or birth-related injury. [...] One of the features of cerebral palsy is the presence of slow, writhing aimless movements.

A study carried out in centers around Europe [15] shows that the prevalence of infants with this condition is about 0.208%. It also shows that the incident rate is much higher for infants born too early and especially for new-borns under 1500 grams. It is of great importance to detect the condition and to treat the infants as early as possible, while the infants experience with abnormal movements still are limited [9].

2.2 General Movement Assessment (GMA)

Infants have a repertoire of distinct movement patterns. One of these patterns emerges and are observable especially when the infants are between 9 and 20 weeks old. The infants' movements are now of a fidgety character, and can be described as an ongoing stream of small, circular, and elegant movements of neck, trunk, and limbs [27]. Abnormal fidgety movements look like normal fidgety movements except that their amplitude, speed and jerkiness are moderately or greatly exaggerated. Absent or abnormal quality of this kind of movements is usually an indicator of a neurological impairment which will manifest itself later in life.

H.F.R. Prechtl et. al. [27] have developed a method for observing the quality of these fidgety movements called General Movement Assessment (GMA). This method has very high specificity (96%) and sensitivity (95%) of its predictions and is therefore a solid basis for the work in this project. The definition of sensitivity and specificity can be found in Section 4.6.1. The method requires training of clinicians which observe the infants and assess the general movements based on a global visual Gestalt perception method as described by Konrad Lorenz [19]. The idea of this method is to assess the infant's movement based the general impression of the whole video, and not by focusing on any details throughout the assessment.

There are however indications that due to the experience needed, GMA is limited in use in ordinary clinical practice because of the importance of maintaining the GMA standards for the clinician, which over time can drift away without sufficient exercise [1, 11]. This has led to the idea of computer-based analysis of general movements as a supplement to verify the GMA results.

2.3 Related work

Much research has been done on the classification and recognition of human movements in the area of video based tracking [24], but very few studies have focused on infants and neurological impairment. A few notable experiments will be mentioned in this chapter.

2.3.1 Infant movement assessment

A factor which is common for most of the work in the field of infant movement assessment is that they are not truly non-invasive and requires some sort of markers or sensors on the infant in combination with the video camera(s). It may be argued that this could disturb the infant and make it more difficult to capture the fidgety movements.

Meinecke et. al. [22] used reflective markers and seven infrared cameras to capture the movements of infants in 3D. Using quadratic discriminant analysis on an optimalized subset of the captured data, an overall detection rate of 73% was obtained.

Experiments with a model representing the body of an infant together with image segmentation of video data was done by Milan Taticek in his bachelor thesis [32]. Due to the lack of video data no substantial results are available, but the thesis shows a promising approach to a solution. Anomalies in the video data can be ignored and removed from the dataset by using the body model, and the resulting video is more robust for further analysis.

ENIGMA (Enhanced interactive general movement assessment) [4] is a software tool developed as a computer-based diagnosis support system for GMA. The purpose of this project was to determine whether it is possible to model objective features from the infant's fidgety movements based on the knowledge of a general movement expert. Motion data in terms of x-, y- and z-coordinates is captured from sensors placed on the infant's ankles, wrists, sternum and forehead. This data is linked with a video recording of the infant. The ENIGMA project was performed at St. Olavs hospital in Trondheim in relation with NTNU. The results from the report states that the project is an appropriate tool for GM experts knowledge elicitation and can provide a basis for detailed discussions on the matter through visual notions.

Another project originating from the data captured here is an experiment regarding the video data and the concept of motiongrams [1]. By extracting the quantity of movement in the image together with the position of the centroid of motion in time and then

calculating the statistical features of these data, the researchers attained a sensitivity of 81.5% and a specificity of 70.0% of diagnosing pasients with Cerebral Palsy.

Harald Kirkerød [18] used the same motiongram data in his term project. The video image was divided into four quadrants with a circle of varying radius in the center. A probability matrix was then generated for how the centroid of motion moved between the separated areas. This matrix was then used as a feature for classification, reaching a maximum result with a sensitivity of 90.0% and a of specificity 86,6% and an overall solid performance. These results were the motivation for pursuing the work performed in this master thesis.

Even though the work done by Andreas Berg [3] in his master thesis is not video based, it is worth mentioning. Using dynamic models for classification of the movements of infants, based on the same motion data as in the ENIGMA project, he reached a result with a specificity of 90.91% and a sensitivity of 85.71%. Parsa Rahmanpour [28] confirms that the fidgety movements in the head and arms are the most discriminating features in his master thesis.

2.3.2 Optical flow in human motion analysis

Optical flow is the term for the velocity field generated by the relative motion between an object and the video camera over time. Further details on optical flow will be provided in chapter 3.3.

Nicolaos B. Karayiannis [16] used optical flow to extract data for detecting neonatal seizures among infants based on video recordings . Motion strength signals produced by regularized optical flow methods are able to capture and quantify the differences between the rapid and jerky movements due to myoclonic seizures and the typically slower and smoother movements not associated with seizures. This method was used in a system for automated detection of the seizures [17], which reached a performance of over 95% for both sensitivity and specificity.

Yacoob and Black [5] uses parameterized optical flow for tracking human motion. This approach assumes an initial segmentation of the body into separate parts. Principle component analysis and linear transformations were then employed to represent and recognize the different activities performed in the videos. The linear transformation made it possible to recognize activities even when time scaling and time shift were encountered.

The Graz University of Technology have done a lot of research on optical flow. Werlberger [33] uses optical flow incorporating prior knowledge of an object's shape for tracking with good results. Paier [26] constructs a framework for controlling video games using only a web camera and his own hands. The most interesting work from this university is however their work with optical flow calculation in real time by delegating the heavy computation of the flow field to a powerful graphics processing unit [34].

2.3.3 Other relevant work

Bobick and Davis [6] introduce temporal templates as a representation of human movement. They construct a binary motion-energy image (MEI) which denotes where motion has occured in the sequence, and then generate a motion-history image (MHI) which represents the recency of a motion in the sequence. Together they form a two component version of a temporal template, which is a static vector-valued image where each point is a function of the motion properties at the corresponding spatial location in an image sequence. Using aerobics exercises as a test domain, these temporal templates are matched against stored instances of views of known actions.

Bregler [8] describes a probabilistic decomposition of human dynamics with different levels of abstractions using mixture models, expectation maximization (EM) algorithm, recursive Kalman and Markov estimation. The technique is demonstrated by classifying human gait categories such as walking, running and skipping in cluttered video sequences, with promising results.

Several studies on movement classification have been performed with success using Hidden Markov models (HMM) [8, 23, 25]. Gaussian Mixture models (GMM) is a method useful for classification of static patterns and HMM for sequential patterns.

2.4 Summary of related work

Optical flow has been used in human motion analysis for numerous projects. The majority of the projects are however trying to identify simple and easily recognized movements like walking, running and waving an arm. This is not the case for the work in this thesis. It is here attempted to separate the normal and abnormal movements of infants in the fidgety age based on a simplified representation of the data that optical flow provides. The principal difference is that most of the other projects are having clear samples of the movements they try to classify for comparison, whereas this thesis only deals with the complex nature of the infants' movements which only an GMA expert can truly assess without any substantial comparison data. One exception is the work of Karayannis [17], but in that project the focus was on the quantity of the motions, and not the quality as is the case here.

3 Materials and theory

The earlier work in this field using the motiongram method [18] produced encouraging results. Combined with the accessibility of an expert in the field, Annette Stahl, this was the motivation for attempting to improve the results using optical flow as a preprocessor for data acquisition for subsequent classification.

This chapter starts with a short description of the dataset used in this thesis. Then a short introduction of the theory behind motiongrams and optical flow are given followed by the motivation for using the optical flow method. Finally, a short introduction to Markov models is given.

3.1 Video data of infants

Video recordings were performed on a study group of 82 infants, of which 15 are confirmed diagnosed with Cerebral Palsy [1]. Each infant was filmed one to five times between the age of 10 to 18 weeks post term for a total of 136[1] recordings, using a stationary video camera (Sony DCR-PC100E) placed above the infant during active wakefulness. The infants were wearing a diaper and a body lying in supine position on a mattress. The recordings were edited by a trained GMA observer (Lars Adde) and optimized for further analysis by omitting all movements due to sensor wire movements or other disrupting movements.

Written consent was obtained from all parents, and the Regional Committee for Medical Research Ethics and Norwegian Social Science Data Services approved the study. The resulting videos are protected by the privacy protection laws of Norway[2], and can therefore only be accessed on a computer without any network connection. The hard disk drive with the videos has to be locked in a safe place when not supervised.

3.2 Theory of motiongrams

The concept of *motiongrams* is based on the waveform displays and spectrograms from the audio domain, from which the features in the audio material can be visualized. This is the idea of Alexander Refsum Jensenius [14], who researched for a method for visualizing longer sequences of movements over time in his Ph.D. thesis.

A motion image is created by calculating the difference between consecutive frames in a video. A motiongram is then made by reducing the $M \times N$ matrix of the pixels of a motion image into $1 \times N$ and $M \times 1$ matrices by calculating the mean value of the colums and rows respectively. Placing these one pixel wide or tall vectors after each

[1] The total number of recordings used in [18] were 137, but the recording 'P063_T03' was not available for this project.

[2] This is why the pictures in this thesis of the infants have been saved in a bitmap format and not a vector format to prevent the possibility of removing the censor mark anonymizing the infants.

other chronologically gives running displays which make it possible to see the location and level of movement in a video sequence over time. Figure 1 from [14] shows the process of creating a motiongram. The data from the motiongrams are then used to calculate the centroid of motion (CoM), which is the position of the centroid of the total movement in a motion image.

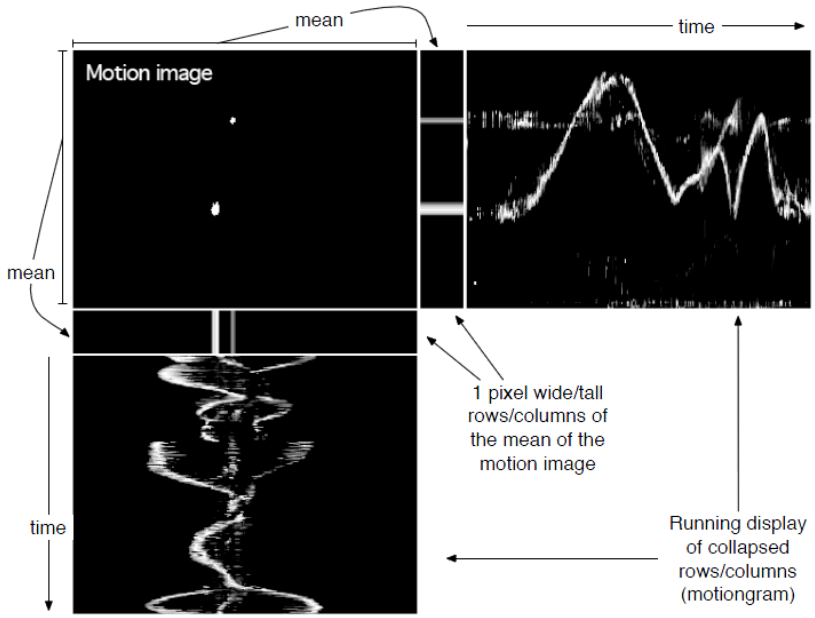


Figure 1: Illustration of the creation of a motiongram. The running display shows the movement of the centroid of motion over time. [14]

3.3 Theory of optical flow

The main source for the following section where a short introduction of optical flow is given, has been the Ph.D. thesis of Annette Stahl [30] together with the book Learning OpenCV [7]. The paper from Tomer Amiaz [2] was used for the description of the coarse to fine method.

Optical flow or image flow is defined as [13]:

> Image flow is the velocity field in the image plane due to the motion of the observer, the motion of objects in the scene, or apparent motion which is a change in the image intensity between frames that mimics object or observer motion.

Stahl [30] defines it further as a 2D velocity field that describes the intensity changes between images. The following paragraphs are cited from Stahl [30] with some minor modifications.

An image sequence is represented by a real valued image intensity function $I(x,t)$ that is continous in space and time. The variable $x = (x_1, x_2)$ denotes the location within a rectangular image domain Ω, and $t \in [0,T]$ labels the corresponding frame at time t.

A link is created between apparent motion and the corresponding intensity variation. The most frequent assumption within these approaches is that the observed intensity I is conserved over time. This means that the intensity at position (x_1, x_2) at time t will be the same as the intensity at time $t + \Delta t$ at position $(x_1 + \Delta x_1, x_2 + \Delta x_2)$ for a small Δt. Using the intensity function $I(x_1, x_2, t)$ along with $u_1(x_1, x_2)$ and $u_2(x_1, x_2)$, which are the two components of the optical flow vector $u = (u_1, u_2)^T$, this results in the equation

$$I(x_1 + u_1\Delta t, x_2 + u_2\Delta t, t + \Delta t) = I(x_1, x_2, t) \tag{1}$$

where $\Delta x_1 = u_1\Delta t$ and $\Delta x_2 = u_2\Delta t$.

Assuming that the brightness varies smoothly over time the term on the left hand side of equation (1) can be approximated by a first-order Taylor expansion at the point (x_1, x_2, t)

$$I(x_1, x_2, t) + \Delta t u_1 \partial_{x_1} I + \Delta t u_2 \partial_{x_2} I + \Delta t \partial_t I + \Theta\left(\Delta x_1^2, \Delta x_2^2, \Delta t^2\right) = I(x_1, x_2, t) \tag{2}$$

Dividing by Δt and dropping the higher-order terms $\Theta(\Delta x_1^2, \Delta x_2^2, \Delta t^2)$ for $\Delta t \to 0$, we obtain the linearised constraint

$$\partial_t I + u \cdot \nabla I = 0. \tag{3}$$

This differential equation is well-known in the literature as the *Brightness Change Constraint Equation (BCCE)* or *Optical Flow Constraint Equation (OFCE)* and was introduced by Horn and Schunck [12]. As the name suggests, Equation (3) represents a

constraint on the optical flow components u_1 and u_2 which requires that the intensity of an object point stays constant along its motion trajectory. Consequently, this equation is a useful approximation for image sequences where the displacements are small, the grey value at a certain position is not influenced by global illumination changes, and where no occlusion occurs in the image sequence.

In order to deal with larger displacements, one can use a hierarchical multiresolution approach which computes the optical flow - using a coarse to fine strategy - at each resolution level. This deals with the problem that, depending on the spatial image frequency, very large motions may cause aliasing along the time frequency axis. The coarse to fine approach estimates the flow in an image pyramid, where the apex is the original image at a coarse scale. The levels beneath it are warped representations of the images based on the flow estimated at the preceding scale. This ensures that the small motion assumption of the optical flow constraint of Equation (3) remains valid.

The downscaling in the coarser images makes sure that movements in the image are between two consecutive pixels, which is a demand to make the algorithm work. It is possible to make the image even coarser to ensure that all movements falls within the two pixel range, but that causes a more heavy computation. If a movement is so large that it surpasses the two pixel range it will cause a faulty flow field for the entire image pair used, as a result of a global optimalization term. In this thesis, four steps of downscaling was used in the algorithm. Its performance will be assessed in Section 4.4.2.

3.3.1 Global and local approaches to optical flow

Global approaches are also known as *dense optical flow* approaches, and this is the case when the optical flow problem is solved over the whole image domain. A velocity is associated with every pixel in an image, calculated from the distance a pixel has moved between two consecutive frames. Horn and Schunck [12] introduced one of the first methods of this approach, and is still regarded as one of the best performers for motion estimation [21].

Local approaches determine the flow at a specific pixel position $x = (x_1, x_2)^T$ by using only the image information in the local neighbourhood. This is also known as *sparse optical flow*, and the main approach is to detect beforehand a subset of points that are to be tracked in the image sequence. The computational cost is much less for the local approach in comparison to the global approach. Lucas and Kanade [20] created a well known algorithm for this approach.

Since the main objective of this thesis is to use optical flow as a preprocessor to produce data that can be compared to the earlier results based on motiongrams, it is natural to choose the global approach in this matter. The global approach will produce an optimized flow field of the total picture, whereas the local approach is more suitable for tracking purposes.

3.3.2 Incorporating Burgers equation in optical flow

While the earlier mentioned theory represents a static viewpoint to image motion processing, Stahl [30] introduces a dynamical one by incorporating Burgers equation as a regularisation term, and summarizes this approach the following way in her thesis:

> Our dynamic image motion approach combines variational motion estimation with motion prediction through a transport process based on the Burgers equation as new physical prior exploiting the knowledge that moving structures should exhibit some inertia, meaning that velocities are not expected to change.

The inertia of a moving structure in the image sequence is calculated from prior images and then predicted for future images in a sliding window manner. A deviation from the prediction motion velocity is penalized in the regularisation term. This results in a more robust tracking of the moving object which is less vulnerable to anomalies in the images or movements. The approach is well illustrated in Figure 2, where you can see the calculated inertia of the moving object expects it to continue to the right. A sudden change of direction makes the velocity flow field continue to the right while the object is going down, but it quickly catches up with the movement in the last frame when it is clear that the movement change was valid.

This approach produces more accurate flow fields than Horn and Schunck and deals with significant noise levels in an improved manner, compared to the traditional approaches. Another advantage with this approach is that it is capable - up to a certain degree - to deal with occlusions. It can also give a robust representation of the rotation of the object. Based on the nature of infants' movements, these are desirable features to look more into for movement classification.

Figure 3 shows a plot of the optical flow field with and without the Burgers equation. The difference is subtle, but the vectors overlaying the object in motion is more exaggerated due to the extra term added by the Burgers equation optimalization.

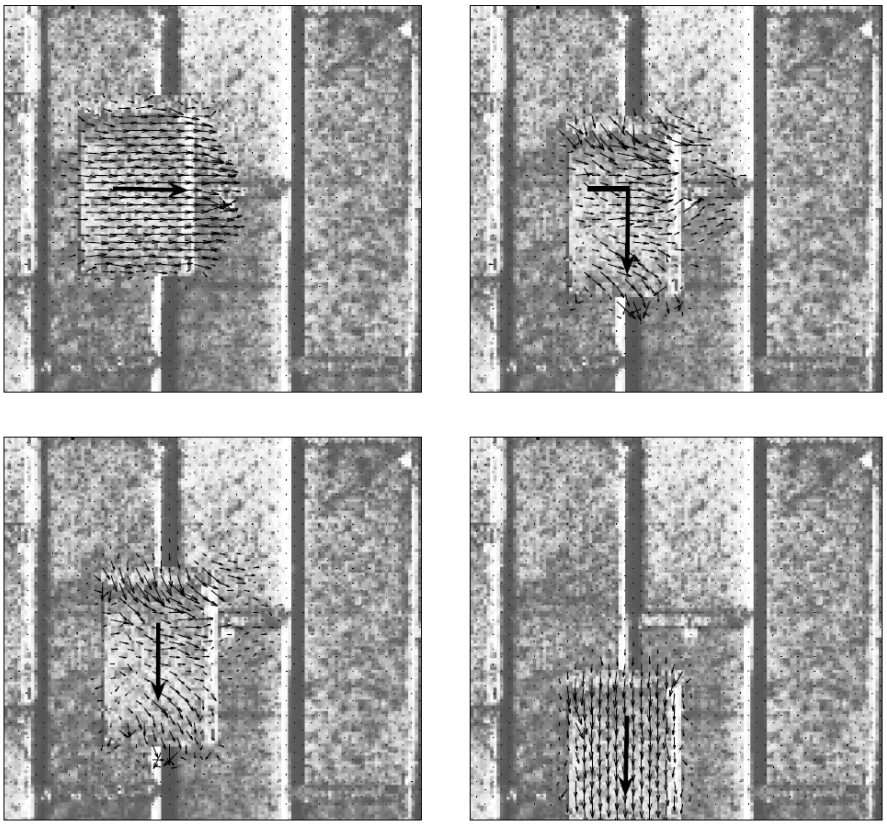
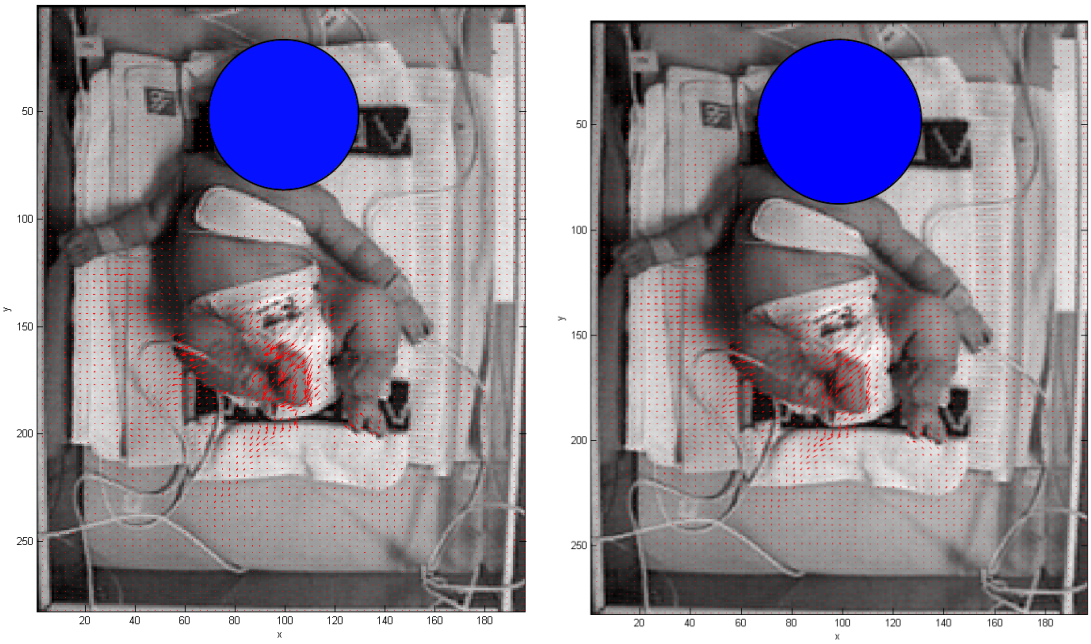


Figure 2: A moving object which abruptly changes the moving direction, as indicated by the large arrows. The figure illustrates the influence of the inertia of a fictive fluid which is trying to track the apparent image motion. [31]



(a) With Burgers equation

(b) Without Burgers equation

Figure 3: A visualization of the optical flow field with and without the Burgers equation.

3.4 Motivation for using optical flow

Both motiongrams and optical flow fields give a representation of the amount of movement from one frame to the next. The difference is that while the motiongram only detects how much the intensity of the pixels in their fixed places have changed, the optical flow field tells where the object with a certain intensity in the image has moved and at what speed. This led to the idea that using optical flow as the data preprocessor instead of motiongrams could provide a more robust and correct calculation of the centroid of motion.

3.5 Theory of Markov models

A *Markov model* (MM) is a statistical model with a number of observable states. The model may change or retain its state according to its probability distribution. The resulting sequence for the Markov model can then be used in classification. Further details on Markov models may be found in the book Pattern Classification [10].

4 Methods

In this chapter, a summary of the different processing steps the videos went through as preparation for the data extraction algorithm is given. How the optical flow fields were generated follows next. Then the methods for extracting the features and filtering the dataset is explained. The chapter ends with details about the simplified version of the Burgers equation and how the classification were done.

4.1 Preprocessing the video data

The videos were prepared for the optical flow algorithm through the following steps.

- Each video was first deinterlaced and each frame was extracted into an image file of the PNG format (Portable Network Graphics).
- Each image were then scaled down from 720 x 576 pixels to 320 x 288 pixels to reduce the computation time for the algorithm. This was assessed to be a reasonable downscaling without losing important image information.
- The images were reduced further by cropping each image to only include the area inside the infant trolley. This was also a solution for placing all the different infants in a satisfactory equal position within the images.
- Lastly, the images were converted into the grayscale PGM format (Portable Graymap Format), since the optical flow algorithm only examines the intensity in the pictures and not the colors.

Figure 4 shows the difference of an image before and after this process. The resulting number of images was about 5000 per recording.

4.2 Generating optical flow fields

There are a lot of optical flow approaches out there, many of which that may be better suited for this matter. The decision of using the chosen algorithm was taken due to the prior knowledge of its robust performance and the convenience of having its creator available for assistance. The algorithm is developed by Annette Stahl [30] for her Ph.D. thesis, and it is based on the Horn and Schunck method [12]. It is a part of a software package licensed by the Ruprecht Karl University of Heidelberg.

After all the video data were preprocessed, the software was installed on two computers and initiated. Since each computer had a dual core processor (Intel Core 2 Duo), two instances of the algorithm were started on each of them. The computation time of optical flow fields took about 2 weeks in total with the designated hardware. This resulted in 923 GB of data, including the preprocessed image files. This was however without the optimalization incorporating Burgers equation. Unfortunately, due to the complex nature of the algorithm using Burgers equation and the time limitation of this project,

Figure 4: To the left is a frame extracted from the video file. To the right is the same frame after the preprocessing, which was used in the algorithm.

it was not possible to compute flow fields using this approach for all the infants in the data set. Our calculations showed that it would have taken 10-12 weeks of computation time with the hardware setup available, so this was skipped. It should be mentioned that the algorithm has not been optimalized in terms of computation time, so it may be possible that it could be improved and with that run much faster.

Optical flow fields with Burgers equation were however computed for four videos, where two of them are of infants diagnosed as healthy and the other two with Cerebral Palsy. This resulted in a total of 85 GB of data.

4.3 Feature extraction from the dataset

The algorithm reads two consecutive images and generates two matrices U and V for the optical flow field, where U consists of the horizontal components of the velocity vector for each pixel, and V consists of the vertical components. When the Burgers equation is included in the algorithm, two additional matrices are generated containing horizontal and vertical vector components for the fictive force field of the movement. These two matrices are then added to the corresponding velocity field matrices to create the robust optical flow field the Stahl principle offers.

4.3.1 Calculation of features

The centroid of motion (CoM) for each video is found in the following way. By first calculating the absolute value of the vectors in the optical flow field, the scalar for each

pixel becomes a positive value for the amount of movement in that particular pixel instead of the horizontal or vertical component of the vector:

$$A = \sqrt{U^2 + V^2} \tag{4}$$

where A is the resulting matrix containing the length of each pixel's velocity vector.

The CoM for each axis, x and y, are found by calculating the mean horizontally and vertically, resulting in one $1 \times n$ vector and one $n \times 1$ vector respectively for the x-axis and the y-axis. The CoM is then found for each direction by using equation 5.

$$\bar{x} = \frac{\sum_{i=1}^{n} m_i x_i}{\sum_{i=1}^{n} m_i} \tag{5}$$

where m_i is the amount of movement and x_i is the pixel position in the vector.

The value of quantity of motion (QoM) for each frame is calculated by taking the total sum of A. The values for the polar coordinates are calculated with the function atan2 in Matlab and equation 6

$$r = \sqrt{(CoM_x - x_c)^2 + (CoM_y - y_c)^2} \tag{6}$$

where (x_c, y_c) is the center of the image.

4.4 Filtering the dataset

An examination of the calculated matrices discovered that also the pixels that did not have any apparent movement were still giving a significant contribution to the optical flow field. Small artifacts due to the image compression, flickering in of the video recording, small variations in the light or tiny movements of the underlay in the trolley can be the reasons for this. The examination showed that these pixels normally had values within the range of 0 to 0.5, while the pixels with actual movement reached values up to 10. Two functions imitating high-pass filters were designed to minimize the values within the range considered as noise, while leaving the higher values as they were. The lower values were only dampened and not removed. This was because of the situations where there are no or little movement by the infant, which still can contain valuable information .

The functions as seen in equations (7) and (8) are designed so that all values below the corresponding threshold values, namely 1 and 2, are dampened by the function. All values above the threshold values remain the same. The output of the functions are illustrated in Figure 5. Both functions were used for the computation of features and it will be specified in the later results which one that has been used.

$$A_{ij1} = \begin{cases} \frac{1}{e^{10(1-Aij)}} & A_{ij} < 1 \\ A & A \geq 1 \end{cases} \tag{7}$$

$$A_{ij2} = \begin{cases} \frac{2}{e^{5(2-A_{ij})}} & A_{ij} < 2 \\ A & A \geq 2 \end{cases} \tag{8}$$

The signals will through the rest of the thesis be denoted with subscripts, e.g. QoM_u for the unfiltered signal, QoM_1 for the signal filtered with 1 as the threshold value and QoM_2 for the signal filtered with a threshold value of 2.

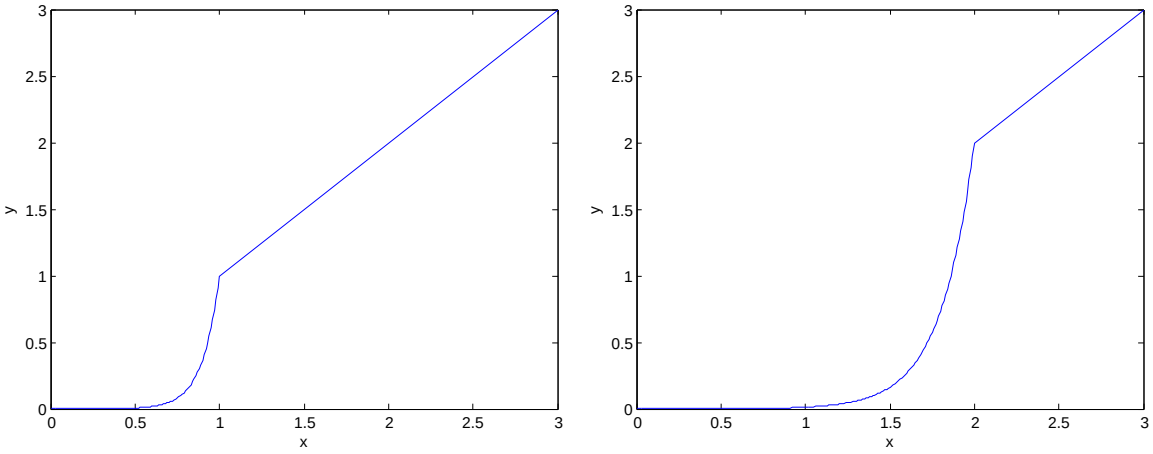


Figure 5: The images show the output produced by the filtering functions, Equation (7) to the left and Equation (8) to the right.

4.4.1 The effect of filtering the data

The result of the filtering of data are illustrated in Figure 6, which is a plot of the QoM through 99 frames of a video. The red line shows QoM_1 which follows the trajectory of QoM_u, the blue line, quite accurately. It is a bit softer around the edges, but overall it removes the noise without disturbing the original behaviour too much. The green line, QoM_2, both removes the noise and dampens the amplitudes quite noticably.

The filtering effect on the CoM is plotted in Figure 7 and shows the section of an image containing the CoM for 99 frames. Filtering the lower range of the signal amplifies the CoM substantially and it is easy to see the similar trajectory of the different signals. The CoM_1 resembles a smooth amplification of CoM_u, while the CoM_2 has a higher amplification, but a more jerky behaviour.

A visualization of the filtering effect can be seen in Figure 8. In Figure 8(a) the contours of the infant are clearly visible together with the earlier mentioned noise. In Figure 8(b), the noise is minimized, but the movement in the lower part of the body and the left arm are preserved. Figure 8(c) has only preserved the larger movement of the infant's legs.

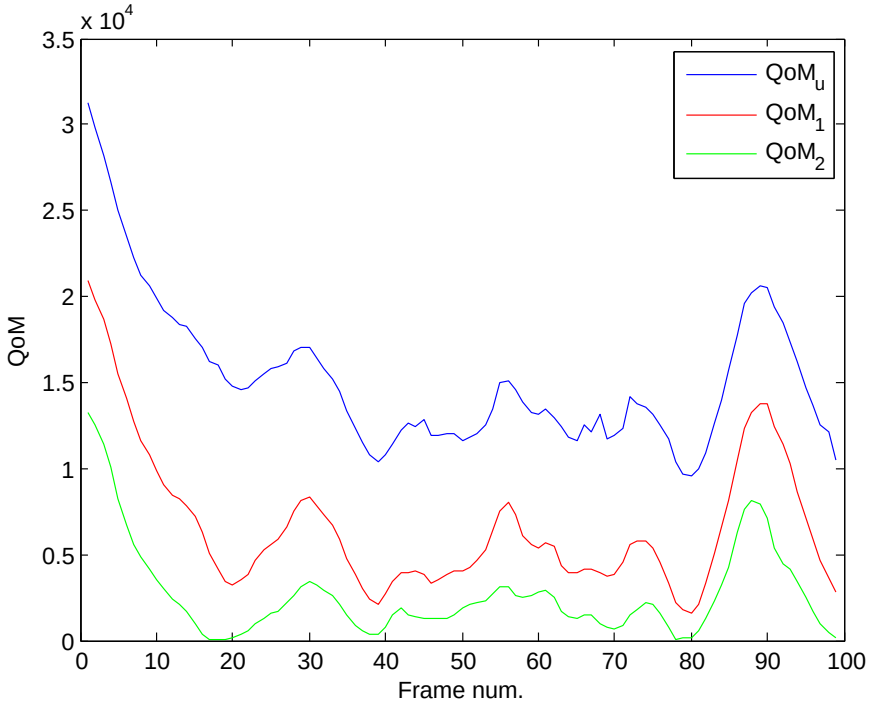


Figure 6: A plot of the QoMs for the different datasets, which clearly shows the differences after the filterings.

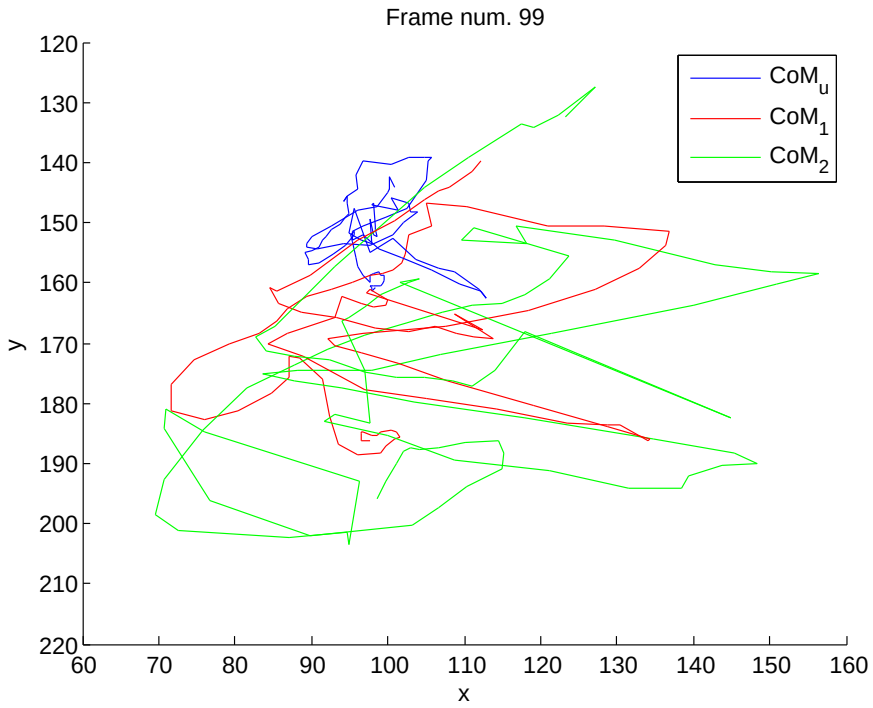
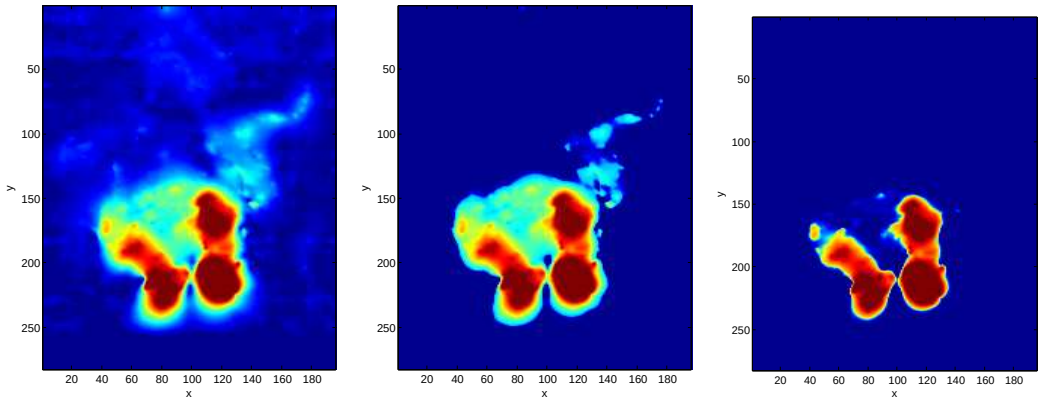


Figure 7: A plot of the CoM for the different datasets. The plot has been cropped around the area containing the CoMs.



(a) $Unfiltered_u$ (b) $Filtered_1$ (c) $Filtered_2$

Figure 8: A visualization of the filtering effect. The colors of the image are the values of the movement in each pixel.

Reason for removal	**Average number of rows removed**
NaN	42.6
Too large QoM	35.4
Only zeros	75.5

Table 1: The average number of rows removed from each patient due to the different faulty results of the algorithm.

4.4.2 Post-processing the computed features

The algorithm does not manage to compute the optical flow fields for absolutely every pair of frames in the videos. One reason is that the coarse to fine method is limited to four steps as a restriction of the computation time. If a too large movement happens from one frame to another which the algorithm cannot detect, it will return a not-a-number value (NaN). Other occurences of faulty optical flow fields have very high values or 0 for the QoM, most likely due to errors with a few frames of the video recording.

These faulty fields had to be removed from the dataset, and it was done after calculating the features of the dataset. First the rows of each recording which contained NaNs, secondly the rows with too large values of QoM and at last the rows of only zeros were removed. The rows with too large values of QoM were detected by checking if the QoM is larger than for an optical flow field consisting of only very high velocity vectors, i.e. a flow field where all values are greater than 10. The average number of rows removed for all the recordings can be seen in Table 1. These numbers were considered as negligible for the overall results.

The features were at last normalized by dividing the features on the height, width and area of each image, resulting in positions between 0 and 1 for the CoM for each picture.

4.5 Simplified version of the Burgers equation

The inviscid Burgers equation is defined as

$$\frac{D}{Dt}u = \frac{\partial u}{\partial v} + (u \cdot \nabla)u = 0, \quad u(x,0) = u_0 \tag{9}$$

Because of the problem with the computation time of the optical flow algorithm incorporating this equation, a simplified version was implemented. The nonlinear term $(u \cdot \nabla)u$, which is the inertia term of the transport process in Equation (9), was omitted. This was the term which increased the complexity of finding a numerical solution of Equation (9). By removing this term the result is a simple temporal differentiation of two consecutive optical flow fields, without the spatial differentiation that the term contributes to.

The feature from this calculation is a simple approximation of the quantity of movement *change* ($QoMC$) between two video frames. It is found by subtracting the horizontal vector component matrix U_i with the former one U_{i-1}, and the same for the vertical vector component matrix V_i and V_{i-1}. The vector length of each pixel in the resulting matrices are then calculated as for the calculation of the CoM in Equation (4). The $QoMC$ is at last found by finding the total sum of the matrix. This was performed for all three datasets.

A plot of the comparison for the different versions of the $QoMC$ can be seen in Figure 9. The simplified versions follow the versions using Burgers equation quite accurately for all three datasets, which means that the simplified version is a decent alternative when there is a demand for a short computation time. The accuracy of the version using Burgers equation is however not reprodusable with the simplified method as illustrated in Figure 10, where the fictive force field from the Burgers equation is plotted on the left versus the simplified version on the right. These illustrations show that while the simplified version represents a satisfactorily quantitative alternative to the advanced version using Burgers equation, the accuracy are not comparable nor usable for a more robust movement tracking purpose.

The idea of exploring features in the limbs and the head of the infant let to the solution of summing up the $QoMC$ for separate parts of the image. A high value for one of the areas would then represent a large change in the movement of a limb or the head, e.g. a kick with the left leg. Five areas were defined as illustrated in Figure 11. The size of each area is calculated relative to the size of the image, since there were minor differences of the cropped images for the different videos.

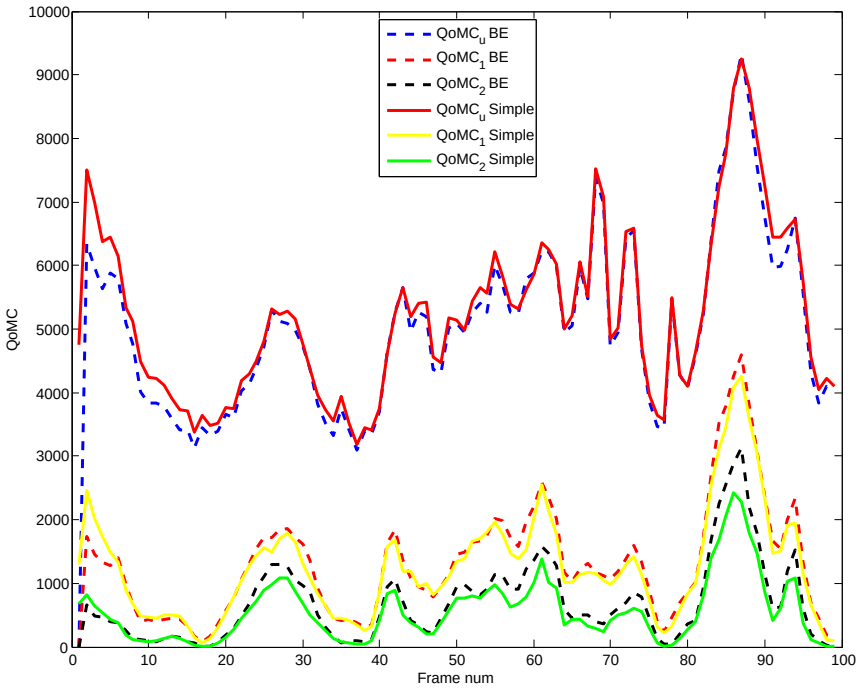


Figure 9: A quantivative analysis of the simplified version versus the Burgers equation version. The $QoMC$s from the simplified version follow the more advanced version quite accurately.

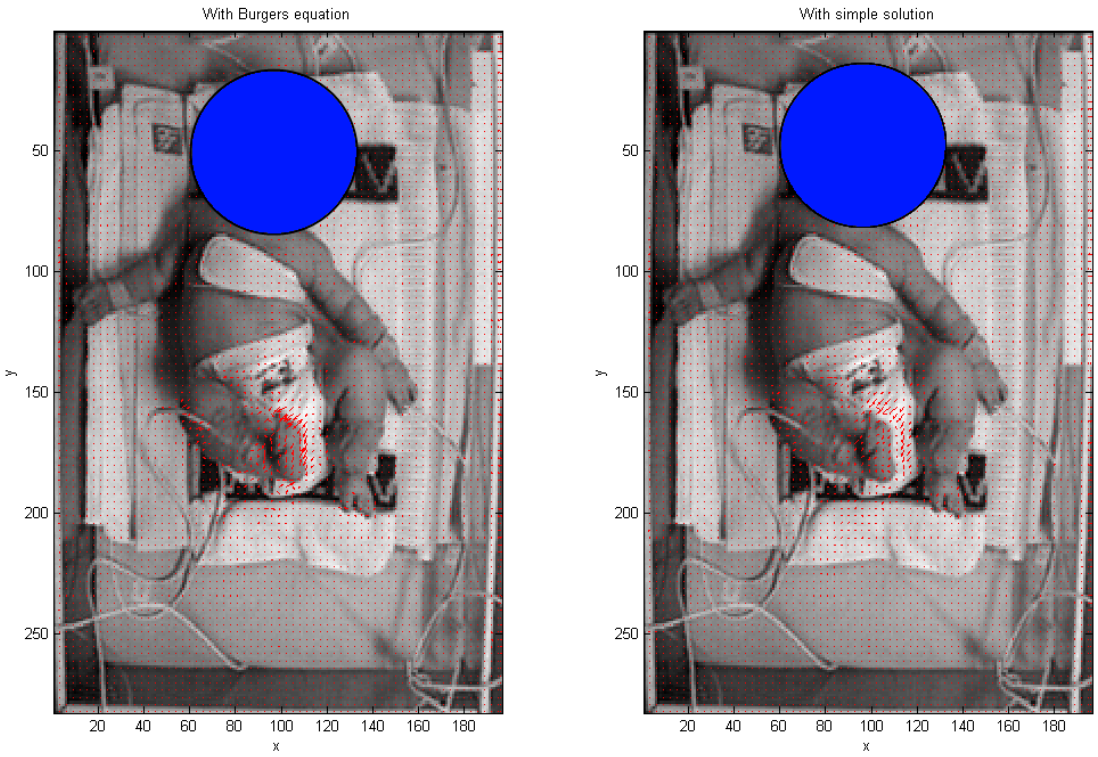


Figure 10: A qualitative illustration of the simplified version versus the Burgers equation version. Here it is clear that the accuracy of the simplified version are not up to par with the advanced version.

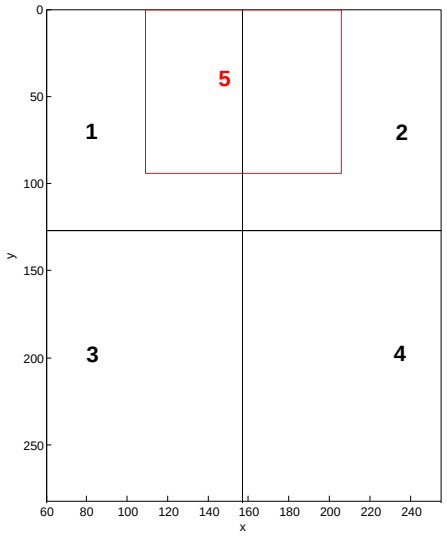


Figure 11: The areas of the division of the image for each infant.

4.6 Classification

Since the main purpose of this project is to compare the optical flow method against the motiongram method, a short recapitulation of the classification definitions and methods are provided here. The following sections are therefore cited from Kirkerød [18] with minor modifications.

4.6.1 Definition of sensitivity, specificity and Youden's index

Sensitivity (SE) and *specificity* (SP) are both indicators of how well a binary classificator is performing. The sensitivity yields the probability for a positive result, given that the condition is indeed positive. The specificity yields the probability for a negative result, given that the condition is indeed negative. In this case, a positive condition means that the patient is diagnosed with Cerebral Palsy. The sensitivity is the most important result for this project, since it indicates that the patient probably has the diagnose CP.

During classification there are 4 situations that can occur:

1. True Positive (TP): Positive condition and positive result
2. False Negative (FN): Positive condition and negative result
3. False Positive (FP): Negative condition and positive result
4. True Negative (TN): Negative condition and negative result

The calculations of these measurments are performed as follows:

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{10}$$

$$\text{Specificity} = \frac{TN}{TN + FP} \tag{11}$$

Youden's index is a rating of the performance of a diagnostic test. The result is a good indicator, but it does not reveal problems in the sensitivity or the specificity. The maximum value is 1 and minimum is -1.

$$\text{Youden's index} = \text{Sensitivity} + \text{Spesificity} - 1 \tag{12}$$

4.6.2 How the classification was implemented

Classification was performed with the function classify in the Statistics toolbox of Matlab. This function allows specification of the type of discriminant function for classification, and all five of which have been evaluated where possible. The five different discriminants are:

1. Linear (LD)

2. Linear with diagonal covariance matrix estimate (LDC)
3. Quadratic (QD)
4. Quadratic with diagonal covariance matrix estimate (QDC)
5. Mahalanobis distances with stratified covariance estimates (MDD)

The abbreviations in parenteses behind each discriminant is how they will be identified throughout the report.

4.6.3 Classification based on Markov models and the normalized matrix

The following section is cited with some modifications from the term project of Kirkerød [18]. This is to be considered a short explanation of the technique used to generate the results using motiongrams that this thesis tries to improve using optical flow. Closer details on this subject is found in his report which is available on the enclosed cd, see Appendix C.

The movement of the *CoM* was tracked through different states of the image. A simple geometrical division of the image was performed to form the areas. Four quadrants make up areas 1 to 4 while the center circle is area 5, which are also the states of the Markov model. By classifying on a variable radius of the center circle different results were reached. The radius is given with a value ranging between 0 and 1, where for instance 0.30 is equivalent to 30% of the total image width. This division is illustrated in Figure 12.

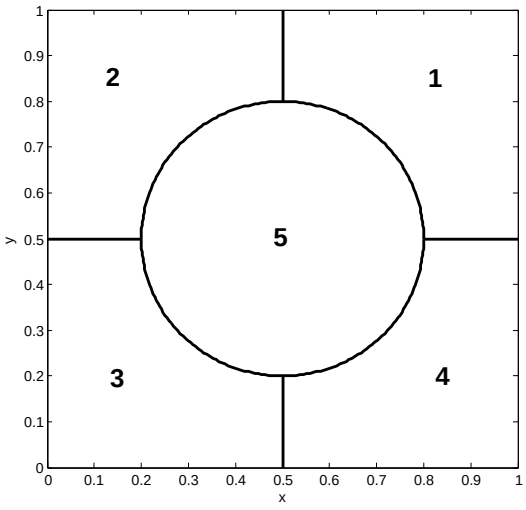


Figure 12: An illustration of the manually created areas. State 5 is the center circle, here with a radius of 0.30.

The probability for a transition changing from one of the n states to another or to itself is defined in a $n \times n$ transition matrix for Markov models. Each row contains the probability for the state i changing to state j, where $i, j = 1 \dots n$. The rows are normalized, which makes the sum of the values in each row equal to 1.

A transition matrix where the total sum of the matrix is equal to 1 is also tested. The total matrix is normalized and will henceforth be referred to as the *normalized matrix* (NM) to separate it from the Markov matrix (MM) . This was done to study the total probabilities of the transitions in the system.

4.6.4 Division of patients for training and testing the classificator

Basically the same training set and test set as in [18] has been used for the classification in the experiments for this thesis. Approximately 2/3 of the set was used for training and 1/3 for testing. It was also made sure that no patient occurs in both groups. Table 2 shows the distribution of the dataset.

Set / Diagnose	Normal	Abnormal	**Total**
Training set	73	16	89
Test set	37	10	47
Total	110	26	**136**

Table 2: The training set and test set for the classification.

5 Results

This chapter begins with a comparison of the results based on motiongrams versus optical flow. The last part includes the new attempts of classification using the features from the simplified version of Burgers equation described in Section 4.5.

5.1 Comparison to term project

Each classification attempt begins with a look at the results based on motiongrams. Then the new results from the optical flow data is presented, first the dataset filtered with a threshold value of 1 ($Data_1$) and second the dataset filtered with a threshold value of 2 ($Data_2$). The unfiltered dataset were left out of the feature extraction algorithm due to low expectation of its performance and is therefore not represented in these results. In retrospect it should maybe have been included for the matter of comparison, but unfortunately time did not suffice.

The entropy and variance of the normalized matrices

Entropy is a measurement for the degree of chaos in a system, see Appendix A for the formula used for its calculation. Classifying with respect to the entropy and variance of the normalized matrix yielded the best result in the motiongram project. As seen in Figure 13, the best result was an SE of 90,0% and an SP of 86,6% at radius 0.22 using QDC as the discriminant function of the classifier. The robustness is however not so reliable with steep inclinations on both sides of the best result.

Using $Data_1$ for the same classification reached an SE of 90,0% and an SP of 73,0% using QD at radius 0.52. Figure 14 shows that the accuracy of the classification is increasing up to radius 0.58, followed by a huge decline. This makes the robustness questionable, even if there is a decent platform for the QD before the decrease.

Figure 15 shows the results using the dataset $Data_2$. The peak of MDD at radius 0.52 is the best result with an SE of 90,0% and an SP of 94,6%, but it is apparent that the steep inclinations on both sides reveals the robustness of this classification. The LD does however sustain a platform between radius 49 and 56 around an average SE of 90% and SP of 70%, but also with steep ramps on both sides.

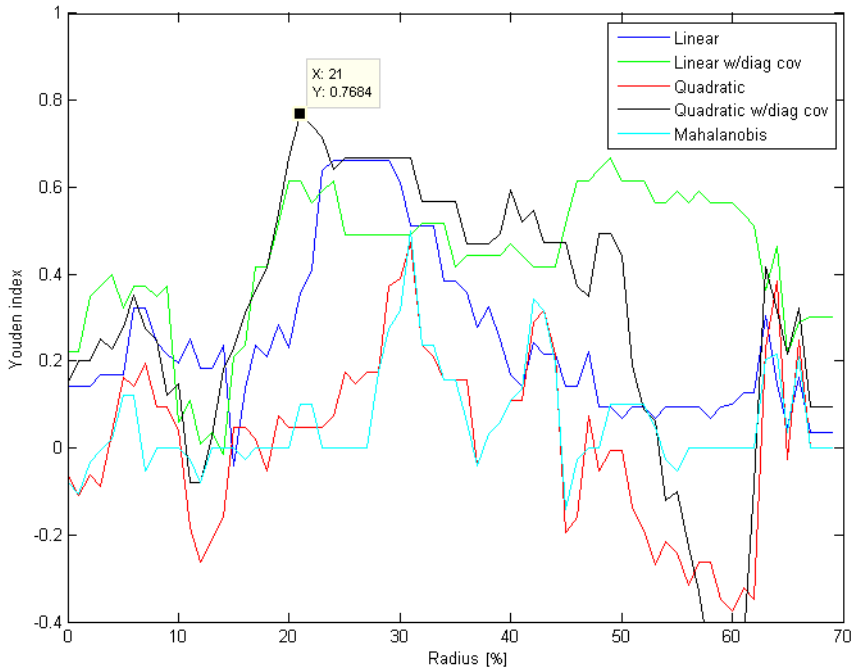


Figure 13: Classification results for the entropy and variance of the normalized matrix from the motiongram project.

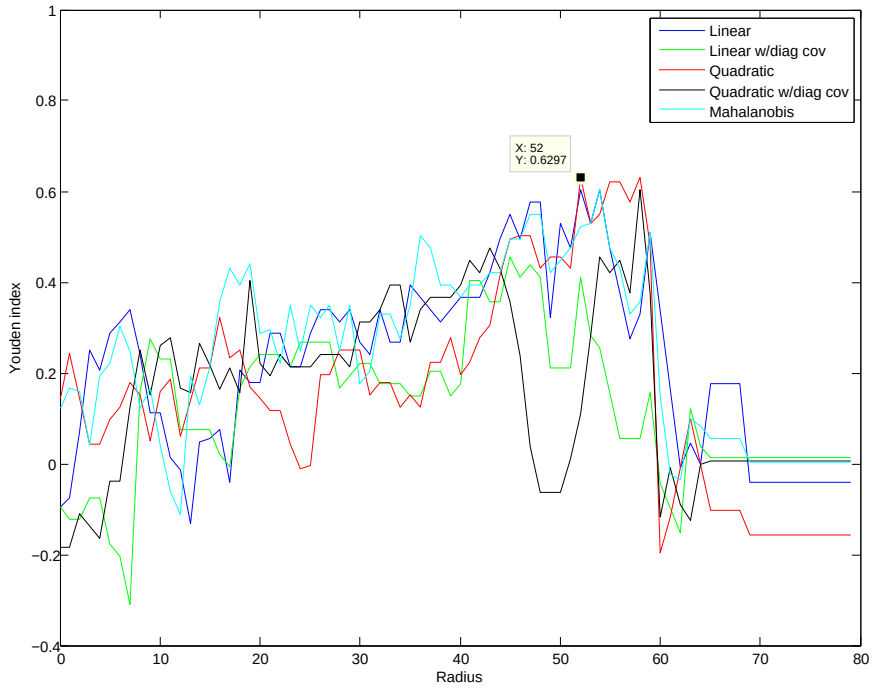


Figure 14: Classification results for the entropy and variance of the normalized matrix using $Data_1$.

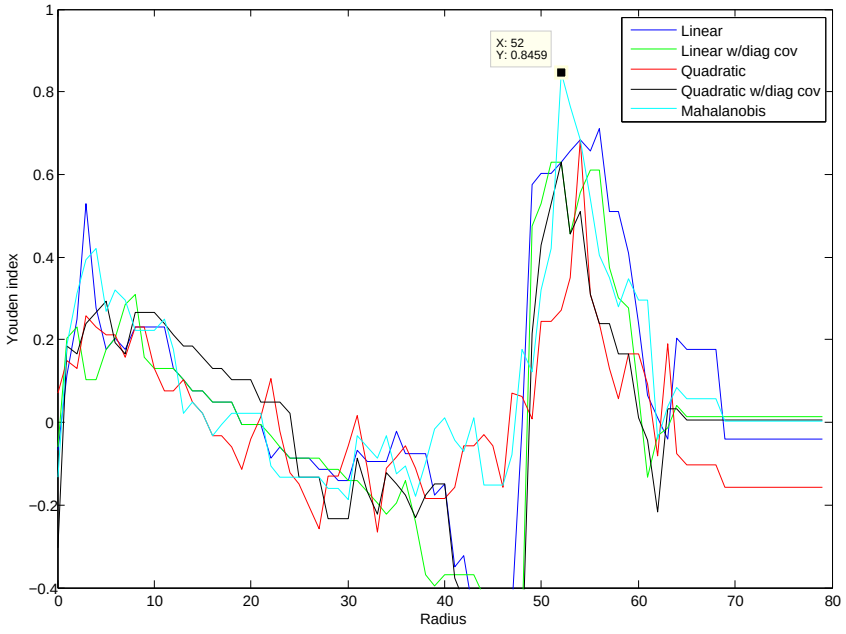


Figure 15: Classification results for the entropy and variance of the normalized matrix using $Data_2$.

The mean and variance of the normalized matrix

Classifying using the mean and variance of the normalized matrix produced a robust result in the motiongram project, as seen in Figure 16. Around the radius 0.28 there is a wide platform where the SE and SP averages about 90% and 70% for the best discriminant function.

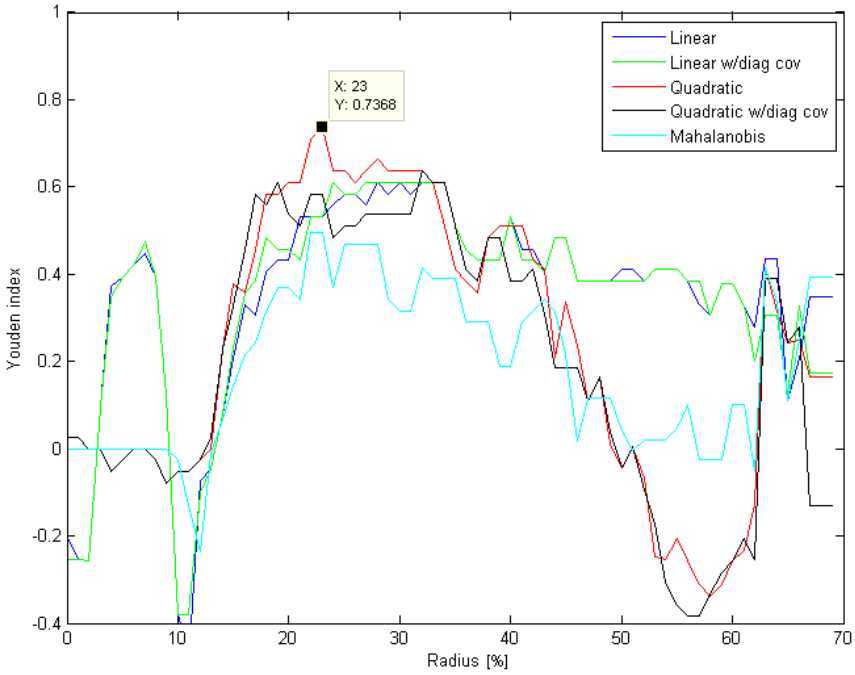


Figure 16: Classification results for the mean and variance of the normalized matrix from the motiongram project.

Figure 17 clearly shows that the dataset $Data_1$ is not close to reproduce the robust results of this classification as for the motiongram project. The peak of QDC at radius 0.54 does produce an SE of 80,0% and an SP of 78,4%, but should be considered as a wild point when studying the surrounding values.

The result of the classification using dataset $Data_2$ is quite similar to the former one as seen in Figure 18, with a large peak at radius 0.54 with an SE of 90,0% and an SP of 81,1% using LDC. Unfortunately, it does also have the steep ramps as the former, making it a weak result.

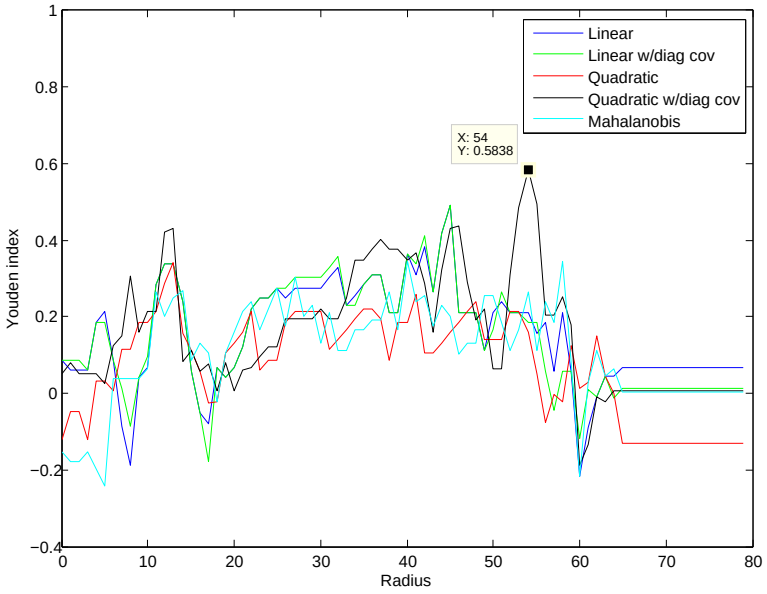


Figure 17: Classification results for the mean and variance of the normalized matrix using $Data_1$.

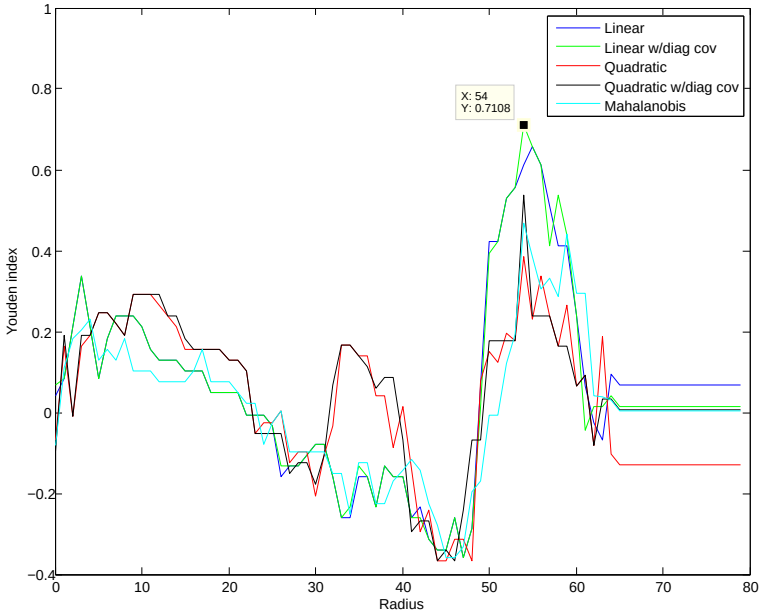


Figure 18: Classification results for the mean and variance of the normalized matrix using $Data_2$.

The entropy of the normalized matrix

For the motiongram project, the classification based solely on the entropy of the normalization matrix produced a decent result as shown in Figure 19. With an SE and SP around 85% and 68% about a wide platform surrounding the radius of 0.30, it is a fairly robust classification of just one single feature. The results using LD and LDC are the same, this is also the case for QD and QDC, which explains why the figures only shows 3 graphs.

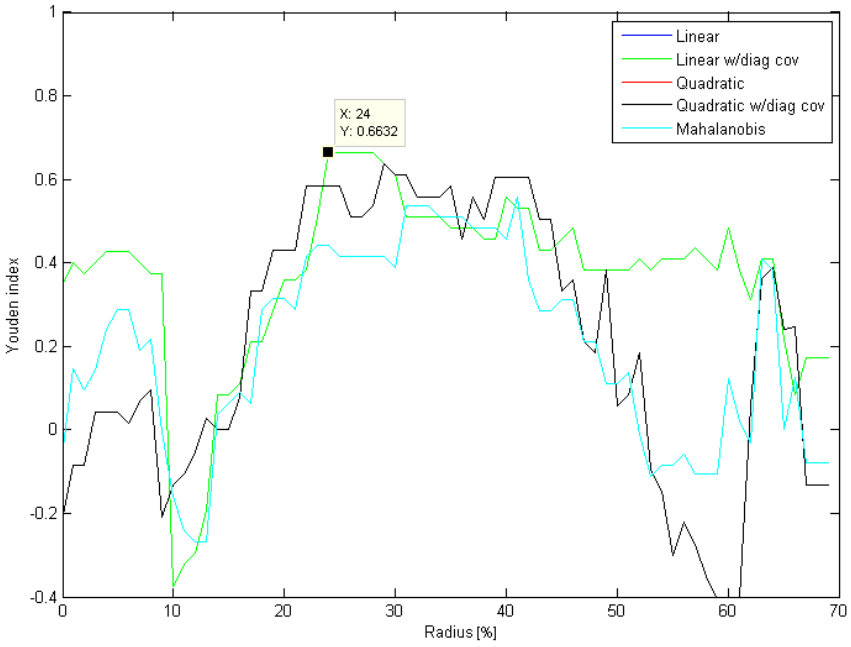


Figure 19: Classification results for the entropy of the normalized matrix from the motiongram project.

The dataset $Data_1$ did not perfom as well as the motiongram classificator as seen in Figure 20, but the dataset $Data_2$ reaches an SE of 80,0% and an SP of 83,8% using LD and LDC for radius 0.56. It is however not more robust than the earlier results as a peak with steep inclinations around the best result, as Figure 21 shows.

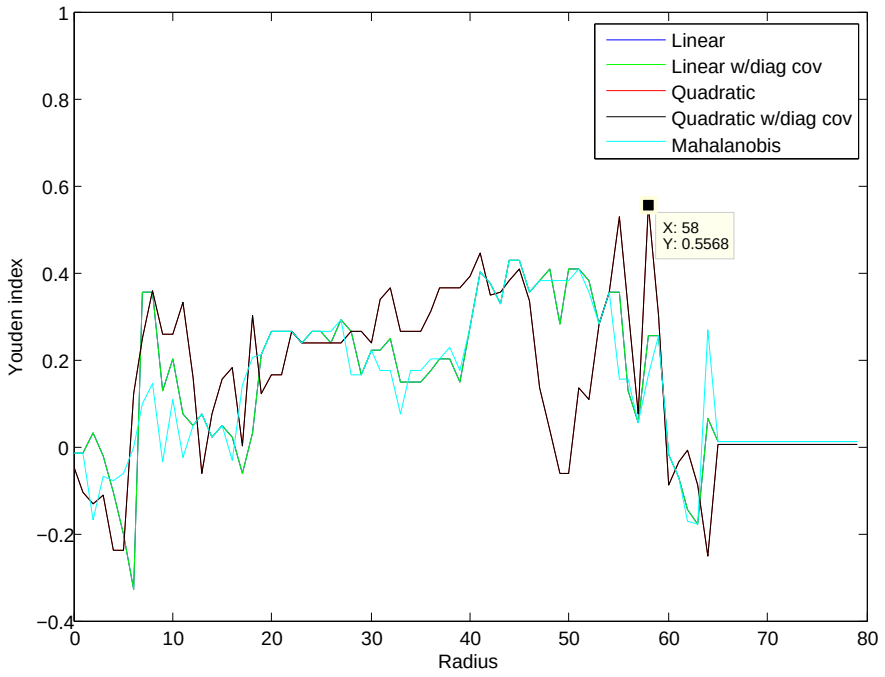


Figure 20: Classification results for the entropy of the normalized matrix using $Data_1$.

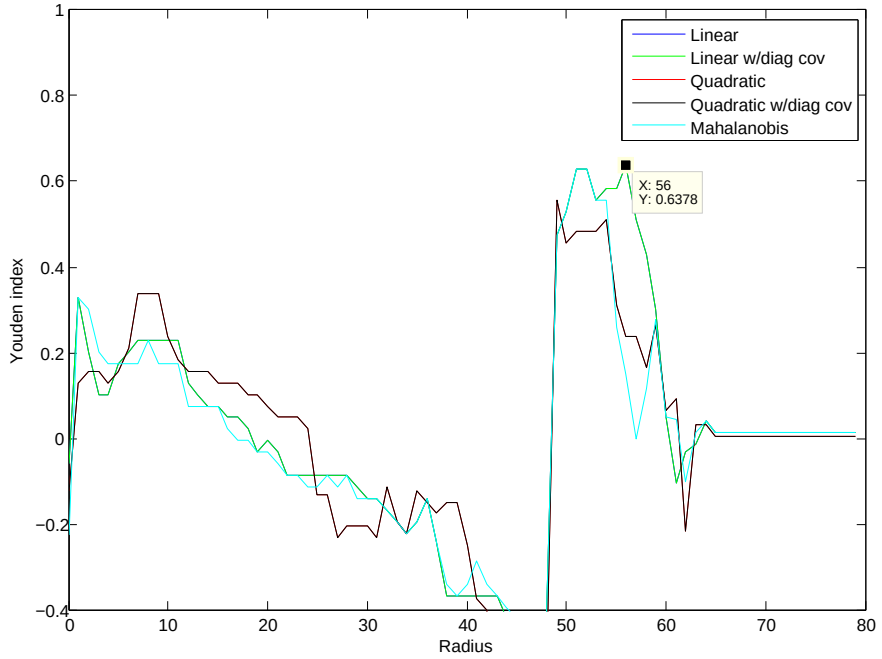


Figure 21: Classification results for the entropy of the normalized matrix using $Data_2$.

The mean and variance of the Markov matrix

Classifying on the real Markov matrix did not produce good results in the motiongram project, as seen in figure 22. The plot has a peak for LDC reaching an SE of 90,0% and an SP of 76,3%, but with steep ramps around it and an overall poor performance.

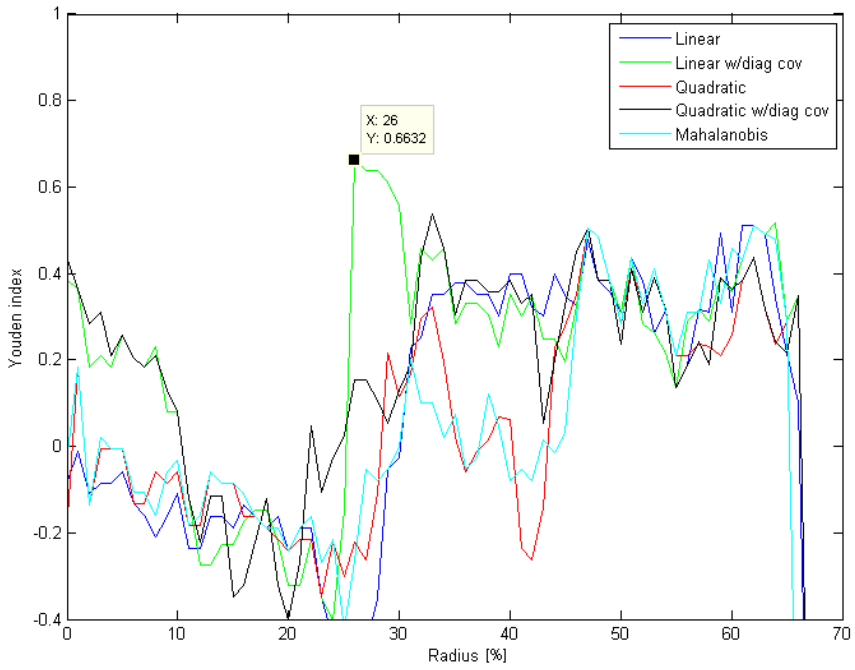


Figure 22: Classification results for the entropy and variance of the Markov matrix from the motiongram project.

Again, the dataset $Data_1$ has a poor performance seen in figure 23, but for $Data_2$ in Figure 24 the performance is decent with a platform between radius 0.52 and 0.57, although it is a bit jagged. The platform has steep inclinations on the sides, but the best result reaches an SE of 100% and an SP of 75,7% using LDC in radius 0.57.

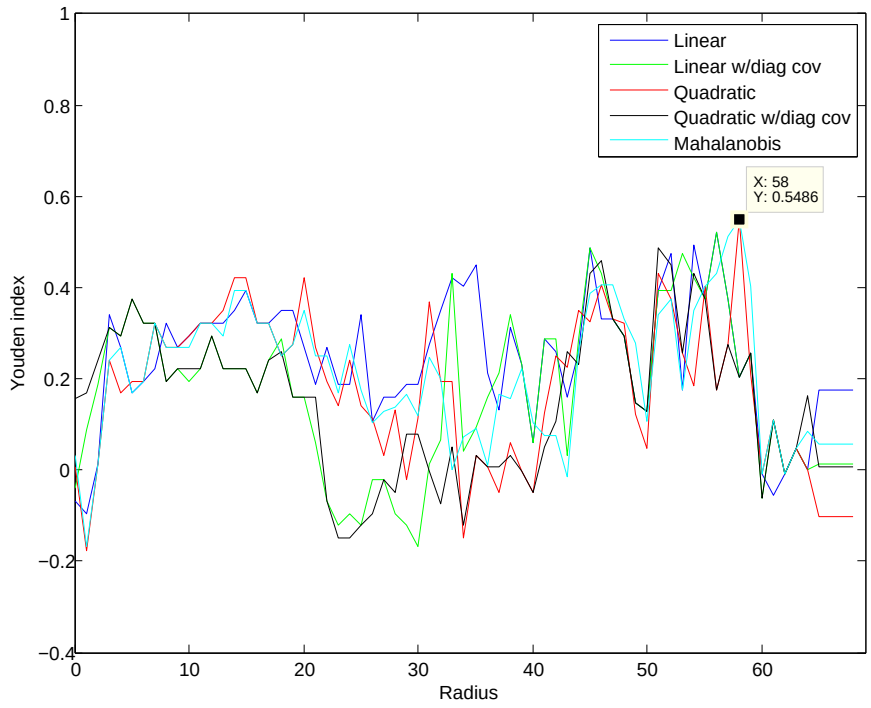


Figure 23: Classification results for the entropy of the normalized matrix using $Data_1$.

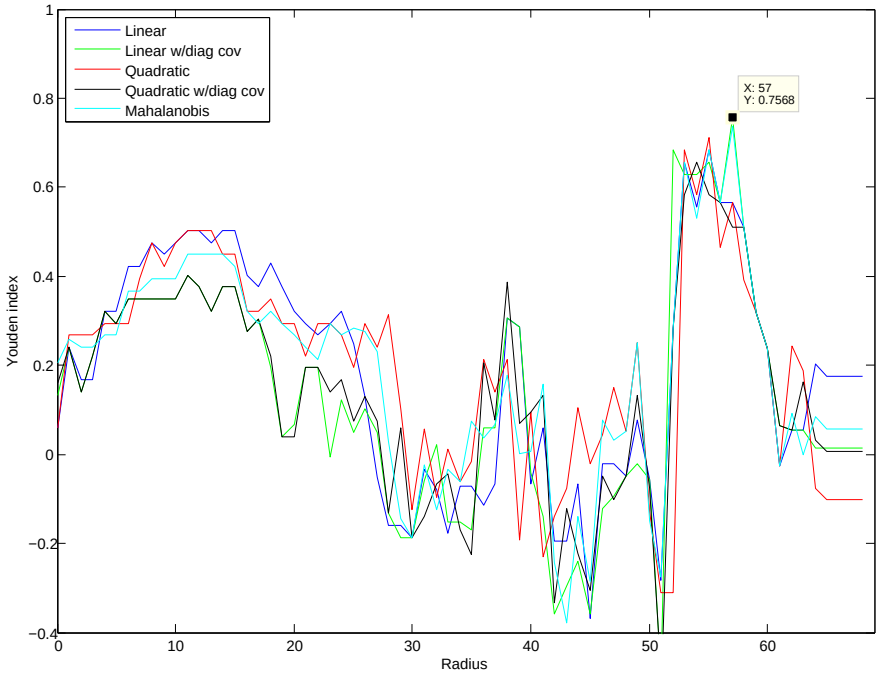
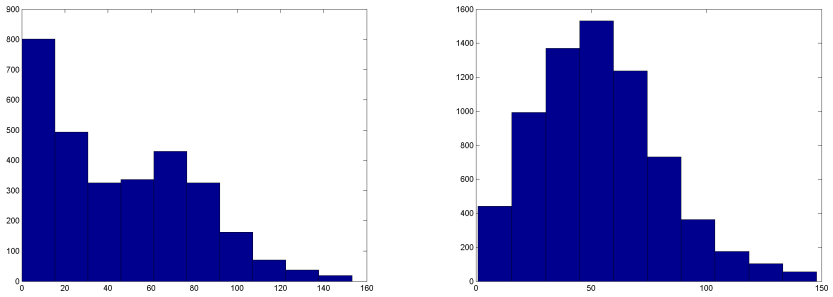


Figure 24: Classification results for the entropy of the normalized matrix using $Data_2$.

5.2 Other results

Evaluating the discrimination of the CoM radius

A few attempts of classifying using other features were attempted. A histogram of ten bars for the distance from the center of the image to the CoMs' positions were plotted for each of the recordings. The idea was that the CoMs for an abnormal infant has either no movement or large movements due to the abnormal fidgety movements. This may result in a histogram with two peaks, one for short radiuses and one for longer radiuses. Normal infants would then have only one peak in the middle as a result of its ongoing stream of movements. Figure 25 gives an example of an abnormal infant on the left and normal infant on the right, which illustrated the idea of this classification. Plots for all the recordings can be found on the enclosed cd, see Appendix C.



(a) $Filtered_1$ (b) $Filtered_2$

Figure 25: A visualization of the filtering effect. The colors of the image are the values of the movement in each pixel.

The classification was then implemented by a binary classification to evaluate the performance of this approach. The datafield of the radius was divided into a variable number of colums ranging from 4 to 20. Classification was performed for each of the values in the range. A plot of results using $Data_1$ as the source is shown in Figure 26, which gave the best result of the two datasets. Through the range from 10 to 14 histogram columns the QD discriminator produces an SE of 80,0% and an SP of 62,2%. The result is not so robust, but it shows that there is a certain discrimination between normal and abnormal infants on this matter.

Classifying on features from the simplified version

Classification using the values of the simplified version for classification was given a few attempts. They did however not produce any results of significance. Classifying the mean and varians of the $QoMC_1$ and $QOMC_2$ with the same classification framework as the other results in this thesis did not manage to discriminate the two groups with

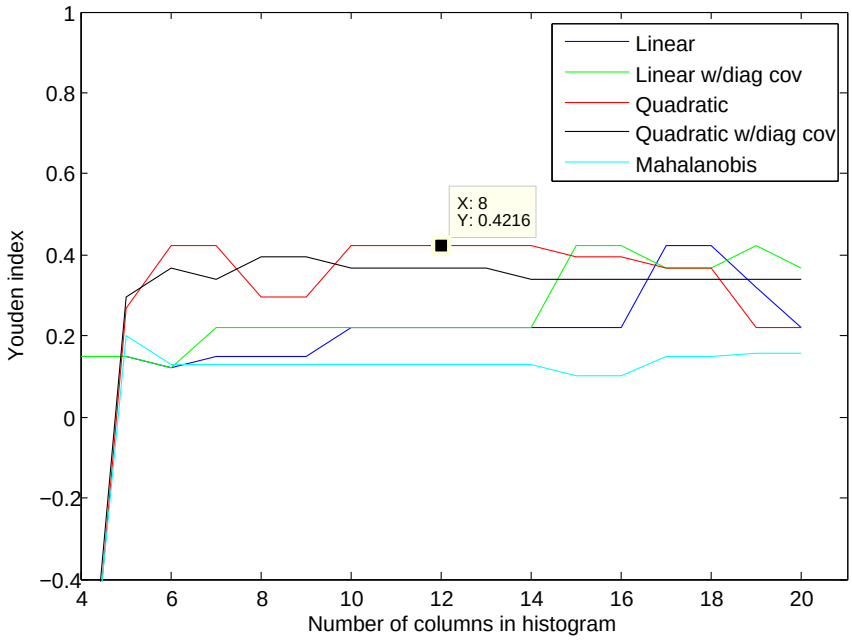


Figure 26: Classification results for the entropy of the normalized matrix using $Data_1$.

an SE of 100% and an SP of 0%, and the time restriction did not allow the construction of a more fitting framework for these features.

5.3 Result tables

Tables 3 and 4 shows the most significant results from the motiongram method and the optical flow method.

Classification features	**SE**	**SP**	**Disc. func.**	**Radius**	**Robustness**
Mean and variance of NM	100,0 %	73,3 %	QD	0,24	High
Entropy of NM	90,0 %	76,3 %	LD and LDC	0,25	Medium
Entropy and variance of the NM	**90,0 %**	**86,6 %**	QDC	0,22	Low
Entropy and variance of the MM	90,0 %	76,3 %	LDC	0,27	Low
Mean and variance of the MM	90,0 %	76,3 %	LDC	0,26	Low

Table 3: A table of the results from the classification using Markov models and the normalized matrix from the motiongram project. NM is short for the normalized matrix, MM for Markov matrix.

Classification features	**SE**	**SP**	**Disc. func.**	**Radius**	**Dataset**	**Robustness**
Mean and variance of NM	90,0 %	81,1 %	LDC	0,54	$Data_2$	Low
Entropy of NM	80,0 %	83,8 %	LD and LDC	0,56	$Data_2$	Medium
Entropy and variance of the NM	**90,0 %**	**94,6 %**	MDD	0,52	$Data_2$	Medium
Entropy and variance of the MM	90,0 %	76,3 %	LDC	0,27	$Data_2$	Low
Mean and variance of the MM	100,0 %	75,7 %	LDC	0,57	$Data_2$	Low

Table 4: A table of the results from the classification using Markov models and the normalized matrix from the optical flow project. NM is short for the normalized matrix, MM for Markov matrix.

6 Discussion

In the previous chapter we saw that the best result based on the optical flow method reached a better result than with the motiongram method, reaching an SE of 90,0% and an SP of 94,6% when classifying . The best result of the motiongram method was an SE of 90,0% and an SP of 86,6%. Both these results should however be treated as *indicators* that there are features of the CoM that can yield a high discrimination between the movement of abnormal and normal infants, and not as a measure of the overall performance of the two methods.

The dataset $Data_2$ reached decent results with an SE of 90% or 80% and an SP of 70% and 80% correspondingly for a platform around the radius of 0.55 for all the four classifications in Section 5.1. This is approximately the same result that was considered as robust in the motiongram project, except that the inclinations enclosing the platforms are much steeper and higher for the optical flow method, as seen in the figures of the results.

The radiuses which provide the best results are on the other hand quite different. For the motiongrams the best results were about radius 0.25, whereas the optical flow method had its best results about radius 0.55. This means that the center area of Figure 12 in Section 4.6.3 covers almost the whole figure, and hence only a small area in each corner constitutes the difference between normal and abnormal movement for the optical flow method. Plotting the CoM for each recording, see Appendix C, does however not reveal any clear distinction between the normal and abnormal movements for the naked eye.

It is worth noting that the dataset $Data_2$ which seemed to have lost some of the movement information during the filtering performed much better than the dataset $Data_1$, which resembled the trajectory of the unfiltered data quite accurately. The performance difference does hovewer seem logical as $Data_2$ will have a quicker and abrupt movement of its CoM than $Data_1$ because of the remaining noise in the latter dataset which contributes to a more soft and sluggish behavior of its CoM.

6.1 Further work

There are still features to explore regarding the nature of the centroid of motion for classification of infant movements. A diamond shaped center area could account for sharper passages to the outer states which could produce a more robust transition matrix. The frequency of the centroid of motion's movement could also yield interesting results. It is however a very simplified representation of the total movement information in an image, so it raises the question of how good discrimination between normal and abnormal infant movement it actually can produce. While it still is an interesting feature worth researching more, it can also be beneficial to consider other approaches to the problem.

The optical flow fields provides several other features which can be worth looking into.

The Burgers equation and to a certain degree the simplified version proposed in this thesis provides a measure of movement change in the picture, which can be considered as a fictive force field. Data for the simplified version, both for the total flow field and the separated areas described in Section 4.5, was calculated and can be found on the enclosed dataset, see Appendix C. The time limitation did unfortunately not allow a proper pursuation of the potential information of these data.

Another feature that could be worth a study is the ability of the optical flow field to detect rotation. If all the vectors overlaying a moving object points in the same direction, it means that the movement is a homogenous translation of the moving object as seen from above. But if the vector field is circular or more chaotic, it could account for a rotation or another complex movement of the moving object. This could be used in a system where first the different movements in a video frame are detected, and then later assessed as complex or simple movements based on the complexity of the overlaying vector fields. Knowing that the normal fidgety movements are considered as small, circular and elegant movements, while the abnormal fidgety movements to a certain extent do not have these features, it could be an interesting discriminator between the two groups.

The algorithm used in this project based on the approach by Horn and Schunck [12] does perform well, but other alternatives should be considered for future projects on this matter. The earlier mentioned project [34] using a GPU for calculating the optical flow fields does also incorporate a new approach based on a total variation regularization term with an L^1 norm, which is demonstrated to be very robust against illumination changes, occlusions and noise. It is also more accurate than the Horn and Schunck approach.

7 Conclusion

Comparing the analysis of infant movements based on optical flow fields to motiongrams culminated in a better maximum result, but with an overall performance which does not clearly separate the two methods. Analyzing how the centroid of motion of the movements in a video clip wanders around in the picture frame is however a decent method to distinguish normal from abnormal movements.

A transition matrix was created by dividing the picture frame into five areas and then recording how the centroid of motion in the video wanders between them. Classifying the patients using the entropy and variance of this matrix gave a sensitivity of 90,0% and a specificity of 94,6%, which is better compared to the motiongram method's result with a sensitivity of 90,0% and a specificity of 86,6%. These are the best results of the two methods, both of which has an overall performance closer to a sensitivity of 90% and a specificity of 70%.

The clinical GMA method has better results then what have been reprodused by analyzing the centroid of motion, but it can still be valid as an objective tool in the diagnostic process. The information in the optical flow fields does however provide complex movement information which are ignored in the process of generating the centroid of motion. Even if there still are features of the centroid of motion which are yet to be explored, it is a very simplified representation of the total movement in an image. A more advanced approach exploiting the movement information of the optical flow field could therefore provide an even better result than what was obtained in this thesis.

References

[1] Lars Adde, Jorunn L. Helbostad, Alexander Refsum Jensenius, Gunnar Taraldsen, and Ragnhild Støen. Using computer-based video analysis in the study of fidgety movements. *Early Human Development*, 85(9):541 – 547, 2009.

[2] Tomer Amiaz, Eyal Lubetzky, and Nahum Kiryati. Coarse to over-fine optical flow estimation. *Pattern Recogn.*, 40(9):2496–2503, 2007.

[3] Andreas Berg. Modellbasert klassifisering av spedbarns bevegelser. Master's thesis, NTNU, Norway, 2008.

[4] P.R. Berge, L. Adde, G. Espinosa, and Ø. Stavdahl. ENIGMA - – Enhanced interactive general movement assessment. *Expert Systems with Applications*, 34:2664–2672, 2008.

[5] Michael J. Black, Yaser Yacoob, and Shanon X. Ju. Recognizing human motion using parameterized models of optical flow, 1997.

[6] Aaron F. Bobick and James W. Davis. The recognition of human movement using temporal templates. *IEEE Trans. Pattern Anal. Mach. Intell.*, 23(3):257–267, 2001.

[7] Gary Bradski and Adrian Kaehler. *Learning OpenCV*. O'Reilly Media Inc., 2008.

[8] C. Bregler. Learning and recognizing human dynamics in video sequences. *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference*, pages 568–574, 1997.

[9] Lidija Dimitrijevìc and Bosanka Jocìc Jakubi. The importance of early diagnosis and early physical treatment of cerebral palsy. *Medicine and Biology*, 12:119–122, 2005.

[10] R.O. Duda, P.E. Hart, and D.G. Stork. *Pattern Classification 2nd edition*. 2001.

[11] JM. Garcia, JLD Gherpelli, and CR Leone. The role of spontaneous genereal movement assessment in the neurological outcome of celebral lesions in preterm infants. *Jornal de Pediatria*, 4:296Ű304, 2004.

[12] Berthold K. P. Horn and Brian G. Schunck. Determining optical flow. *Artificial Intelligence*, 17:185–203, 1981.

[13] R. Jain, R. Kasturi, and B. G. Schunck. *Machine Vision*. McGraw-Hill, Inc., 1995.

[14] Alexander Refsum Jensenius. ACTION –SOUND – Developing methods and tools to study music-related body movement. *Ph.D. thesis*, 2007.

[15] Ann Johnson. Prevalence and characteristics of children with cerebral palsy in europe. *Developmental Medicine & Child Neurology*, 44(09):633–640, 2002.

[16] Nicolaos B. Karayiannis, Senior Member, Bindu Varughese, Guozhi Tao, James D. Frost, Merrill S. Wise, and Eli M. Mizrahi. Quantifying motion in video recordings of neonatal seizures by regularized optical flow methods. *IEEE Transactions on Image Processing*, 14:890–903, 2005.

[17] Nicolaos B. Karayiannis, Guozhi Tao, James D. Frost Jr., Merrill S. Wise, Richard A. Hrachovy, and Eli M. Mizrahi. Automated detection of videotaped neonatal seizures based on motion segmentation methods. *Clinical Neurophysiology*, 117:1585–1594, 2006.

[18] Harald Kirkerød. Motiongram classification for infant movement assessment. *Term project*, 2009.

[19] Konrad Lorenz. Gestalt perception as a source of scientific knowledge. english translation from a german paper in 1959. *Studies in animal and human behaviour, vol. II*, pages 281–322, 1971.

[20] Bruce D. Lucas and Takeo Kanade. An iterative image registration technique with an application to stereo vision. In *IJCAI'81: Proceedings of the 7th international joint conference on Artificial intelligence*, pages 674–679, 1981.

[21] B. McCane, K. Novins, D. Crannitch, and B. Galvin. On benchmarking optical flow. *Comput. Vis. Image Underst.*, 84(1):126–143, 2001.

[22] L. Meinecke, N. Breitbach-Faller, C. Bartz, R. Damen, G. Rau, and C. Disselhorst-Klug. Movement analysis in the early detection of newborns at risk for developing spasticity due to infantile cerebral palsy. *Human Movement Science*, 25:125Ű144, 2006.

[23] Dorthe Meyer. Human gait classification based on hidden markov models. In *3D Image Analysis and Synthesis '97*, pages 139–146, 1997.

[24] Thomas B. Moeslund, Adrian Hilton, and Volker Krüger. A survey of advances in vision-based human motion capture and analysis. *Computer Vision and Image Understanding*, 104:90Ű126, 2006.

[25] Carlos Morimoto, Yaser Yacoob, and Larry Davis. Recognition of head gestures using hidden markov models. In *In Proceeding of ICPR*, pages 461–465, 1996.

[26] Wolfgang Paier. Flowball 2009, 2009. Bachelor Thesis.

[27] Heinz F R Prechtl, Christa Einspieler, Giovanni Cioni, Arend F Bos, Fabrizio Ferrari, and Dieter Sontheimer. An early marker for neurological deficits after perinatal brain lesions. *The Lancet*, 349:1361–1363, 1997.

[28] Parsa Rahmanpour. Features for movement based prediction of cerebral palsy. Master's thesis, NTNU, Norway, 2009.

[29] Rod Seelay, Trent Stephens, and Phil Tate. *Essentials of Anatomy & Psysiology.* McGraw-Hill Science/Engineering/Math, sixth edition, 2006.

[30] Annette Stahl. *Dynamic Variational Motion Estimation and Video Inpainting with Physical Priors.* PhD thesis, Ruprecht-Karls-Universität Heidelberg, 2009.

[31] Annette Stahl, Paul Ruhnau, and Christoph Schnörr. A distributed-parameter approach to dynamic image motion, 2006.

[32] Milan Taticek. Classification of cerebral palsy based on video sequence evaluation. *B.Sc. Thesis*, 2007.

[33] Manuel Werlberger. Globally optimal tv-l1 shape prior segmentation. Master's thesis, TU Graz, Austria, 2008.

[34] C. Zach, T. Pock, and H. Bischof. A duality based approach for realtime tv-l1 optical flow. In *Pattern Recognition (Proc. DAGM)*, pages 214–223, Heidelberg, Germany, 2007.

A The theory of entropy

Entropy is an expression for the degree of chaos in a system. The formula for calculating the entropy is:

$$H(X) = -\sum_{n} P(x) log_2[P(x)] \tag{13}$$

Here $P(x)$ denotes the probability that X is in state x.

B Software

List of software with version numbers used in this project:

1. MATLAB v. 7.6.0.324 (R2008a)
2. Mplayer v. 1.0 RC2 4.2.3
3. Image Magick v. 6.5.7

C Contents of the CD

The plots of each recording's CoM can be found in the folders CoMplot_dataset1 and CoMplot_dataset2.
The dataset used in this project is located in the file workData.mat.
All the Matlab code for this project is located in the folder code.

The report and working data from the motiongram project [18] is located in the folder motiongram.

D The patient IDs with corresponding diagnose

Patient ID	GMA outcome	Patient ID	GMA outcome
1	Normal	46	Normal
2	Normal	52	Normal
3	Normal	54	**Abnormal**
5	Normal	56	Normal
8	Normal	57	Normal
9	Normal	58	Normal
10	Normal	60	Normal
11	Normal	61	Normal
12	Normal	62	Normal
13	Normal	63	Normal
14	Normal	64	**Abnormal**
15	Normal	66	Normal
16	Normal	67	**Abnormal**
17	Normal	68	Normal
18	Normal	69	**Abnormal**
19	Normal	71	**Abnormal**
20	Normal	72	Normal
21	Normal	74	Normal
22	Normal	75	Normal
23	Normal	79	Normal
24	Normal	80	Normal
25	Normal	81	Normal
26	Normal	83	Normal
27	**Abnormal**	84	Normal
28	Normal	85	**Abnormal**
29	Normal	87	Normal
30	Normal	88	Normal
31	Normal	89	Normal
32	Normal	90	Normal
34	Normal	93	Normal
35	Normal	94	Normal
36	Normal	96	**Abnormal**
37	Normal	99	Normal
38	Normal	100	Normal
39	Normal	101	Normal
40	**Abnormal**	102	Normal
41	Normal	103	**Abnormal**
42	Normal	104	Normal
43	Normal	108	**Abnormal**
44	**Abnormal**	111	**Abnormal**
45	**Abnormal**	112	**Abnormal**