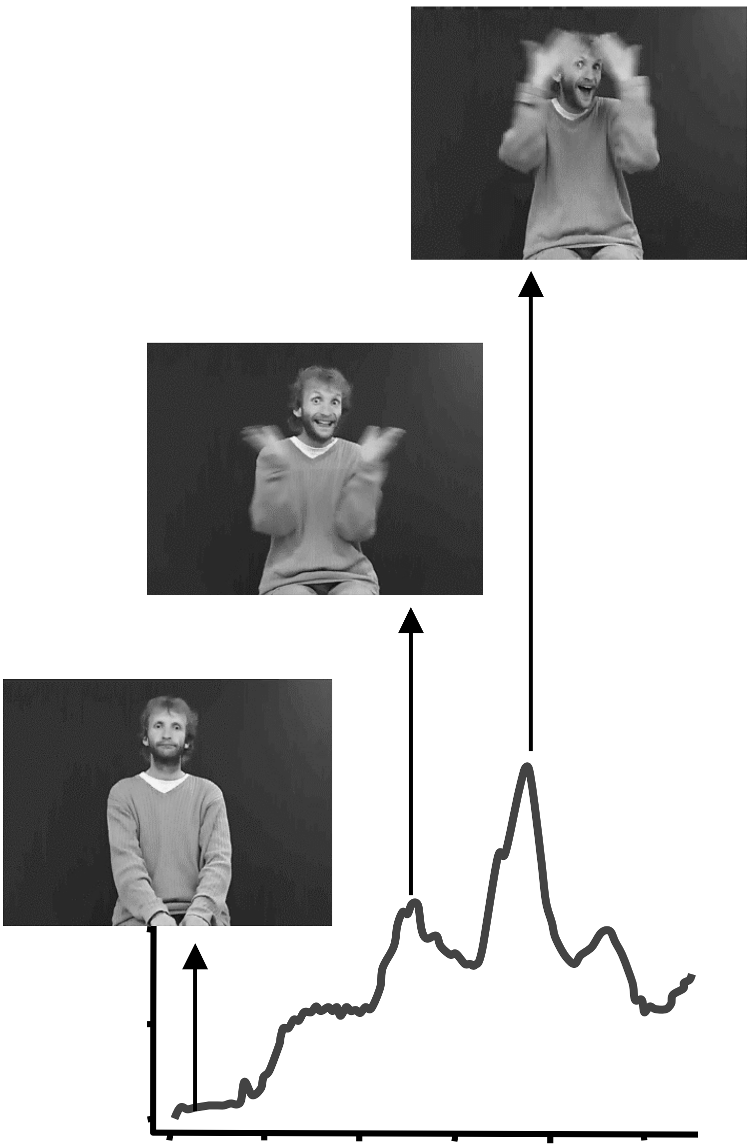
Video Analysis of Sign Language



SAMBA/32/00

Line Eikvil

December 2000

Norsk Regnesentral
ANVENDT DATAFORSKNING

NR-notat/NR Note

Tittel/Title:
Video Analysis of Sign Language

Dato/Date: December
År/Year: 2000
Notat nr: SAMBA/32/00
Note no:

Forfatter/Author:
Line Eikvil

Sammendrag/Abstract:
In this report we have described an initial study on the analysis of sign language videos based on image analysis. The purpose of this study has been to perform a preliminary study of potential approaches to the problem of detecting breaks and extracting highlights from sign language videos.

The reason for looking into this problem is that Møller Kompetansesenter in Trondheim, which is a centre that develops educational tools and materials for deaf people, wants to develop a sign language processor with a flexibility and functionality comparable to that of word processors. In addition to basic video editing tools, they want such a system to offer higher-level tools that can produce an overview of the contents by automatically extracting breaks and highlights.

There are two problems to be solved; detection of breaks and extraction of key frames. In this study we have concentrated on the first part of the problem, while some initial experiments have been performed with the second problem.

Experiments with the developed methods have given good results for break detection. The classification is however based on hard thresholds and some errors are unavoidable. Even better results could possibly be obtained using a statistical classification method. For the key frame extraction, we need sign language expertise to evaluate the results, but initial feedback on the results indicate that they seem interesting.

Emneord/Keywords:	Sign language analysis, Break detection, Key-frame extraction.
Tilgjengelighet/Availability:	Open
Prosjektnr./Project no.:	220070
Satsningsfelt/Research field:	Image and video analysis
Antall sider/No. of pages:	35

Norsk Regnesentral / Norwegian Computing Center
Gaustadalléen 23, Postboks 114 Blindern, 0314 Oslo, Norway
Telefon 22 85 25 00, telefax 22 69 76 60

Video Analysis of Sign Language

Line Eikvil

December 2000

Contents

Chapter 1

Introduction

In several countries sign language is now accepted as the first language for deaf people. Sign language does not have a general written form and documents are therefore produced using live video and video-tapes. Deaf students will for instance make their exams, tests, homework etc. on video cassettes. The flexibility of this medium is however very limited. Without advanced video editing equipment even small changes can only be obtained by reshooting the sequence.

For written language, word processors have made the writing process very flexible compared to old technology based on typewriters. Texts can easily be manipulated with functionality for copying, cutting, pasting and moving text. For sign language, however, the tools are still at the typewriter stage. Hence, new tools for handling sign language are needed. Møller Kompetansesenter in Trondheim is a centre which among other tasks, develops educational tools and materials for deaf people. They would now like to develop a sign language processor for the deaf with a flexibility and functionality comparable to that of word processors.

As a sign language document is a video, basic video editing tools will obviously be needed for such a tool. In addition to this, it would be desirable to offer a few more higher level tools. With video as the medium, it can be difficult to search documents and to get an overview of the contents without playing through the whole video. Hence, tools which can produce an overview of the contents by automatically extracting breaks and highlights, could be very useful. This will require tools based on methods from image and video analysis, which can perform an intelligent analysis of the contents of the video.

The purpose of the work described here has been to perform a preliminary study of potential approaches to get an indication of whether development of such tools is possible. This is a very small study, and most of the problems will not be solved within this project. Our aim is therefore rather to find out through some initial experiments whether this is a problem that can be solved, using appropriate methods. In addition, we want to implement a simple prototype.

Segmentation and extraction of highlights from sign language videos pose a new problem. There exists some work on segmenting and extracting highlights from for instance broadcast news. However, video sequences containing sign language have very different characteristics, and are for instance much more static compared to that of broadcast news. Automatic segmentation of news and other TV-programs are often performed based on the detection of changes between shots and scenes. Video sequences with sign language will however usually contain no cuts and no scene changes, just one person sitting in front of the camera, moving mainly his/hers hands. This makes the problem much more difficult, and closer analysis of the contents are necessary.

Some research has been done on the analysis of sign language videos in order to try to recognise signs [2, 3, 4, 5]. The gesture recognition process, in general, may be divided into two stages: the motion sensing, which extracts useful data from hand motion; and the classification process, which classifies the motion sensing data as gestures. Techniques for tracking hands, e.g. like Pentland and Starner [4] have used for sign language recognition, could probably be useful also for the video segmentation. In this small preliminary study, there was however not enough time to develop and implement the methods needed for this. Also, the system should typically be lightweight and be able to run on a standard PC. This means that computationally complex and time-consuming methods should be avoided. Our approach for the prototype has therefore been to make it simple. If simple methods can bring us towards the goal, it is very probable that these methods can be even further developed and refined to give the results we need.

In this report we will start in Chapter 2 by describing the methods we have chosen to use for the current problem. Chapter 3 contains a documentation of the software prototype which has been developed. In Chapter 4 we will describe some initial experiments performed with these methods and the obtained results. Finally, a summary and a discussion of the work is given in Chapter 5. Details of the software and the listings of the results are given in the appendices.

Chapter 2

Methods

There are two problems to be solved. First we want to detect breaks to split the video up in natural shorter sequences. Then we want to extract key frames from each sequence, which can help to give an overview of the contents of the sequence. In this study we have focused on the first part of the problem, i.e. to detect breaks. This process is described in Section 2.1 of this chapter. However, some initial experiments have also been performed with the second problem in order to get some knowledge about what kind of methods that might work. This initial study is described in Section 2.2.

2.1 Detection of breaks

The initial assumption was that pauses could be detected based on the amount of movement in the image as the hands will be at rest during pauses. However, the hands are also at rest at other times, and even when the hands are at rest, the person might be moving his head or body. Hence, although the amount of movement, i.e. the difference between images, is an important feature, this is not enough to determine whether there is a break.

An additional important characteristic of a break is that when using sign language, the person will rest with his/her hands in the lap to mark a pause. We therefore adjusted our goal to detect the frames where the hands are at rest in the lap. To accomplish this a method to determine at least the approximate position of the hands is needed. We have obtained this by using a simple approach where we, rather than identifying the exact position of the hands, identify a few regions and decide whether the hands are present in these regions.

The detection of breaks is performed through two phases. First, an initialisation phase where the regions of interest are identified based on the hands' position during a break. In this phase parameters specific to the current sequence are also estimated. This initialisation takes place during the first few frames of the video, and the process is described in Section 2.1.1.

In the second phase, the actual analysis takes place. Here, each frame of the video is analysed. The approximate position of the hands is determined based on the regions of interest and the skin-colour model defined during the initialisation. Depending on the assumed position of the hands, each frame is then classified as break or action. The details of this process is described in Section 2.1.2.

2.1.1 Initialisation

The purpose of the initialisation phase is to determine the regions of interest, the skin colour model and estimate parameters needed for classification of frames during the analysis. However, to facilitate the problem, we first needed to make the following assumptions:

1) The background is homogeneous and stationary.

2) The person is sitting in the pause position when the video starts.

3) The hands are brighter than the background when resting in the pause position.

This is not ideal, but we had to make some restrictions to be able to develop the prototype within this very limited study.

The second requirement means that we assume that in the frames during the initialisation phase, the person will sit with his/her hands in the lap. Hence, if we are able to identify the hands during the initialisation we can analyse the same region later, to find if the hands are in the pause position. It can however be difficult to determine whether there are two hands or just one hand in this area as one hand will often be covered by the other during a pause. During the analysis it will therefore not be sufficient to analyse the region covered by the hands in the pause position. The area where the hands move during signing must also be analysed. Both these regions of interest need therefore be determined during initialisation.

Defining regions of interest

The approach for locating the regions of interest starts by finding the position of the person within the frame, and then identifying the two regions corresponding to the upper and lower part of the person. The lower region should then contain the hands when they are in pause-position.

The estimation of the regions is performed as follows. First the edge image is computed. This is an image which gives the strength of the gradients in the image. The reason for using the edge image rather than the original colour image, is that we expect the contour of the person to give strong gradients against the background. This means that to identify the position of the person in the image, it is not necessary to make any assumptions on the colours of the background or the foreground.

It is sufficient to compute the edge image for one band only, and we have used the red band of the colour image. When the edge image has been computed, we threshold it to emphasize the strongest edges. We know that there will be fewer strong edge pixels than background pixels, and we have used a percentile thresholding method. This method will set the specified percentage of the weakest edge-pixels to background and the remaining percentage to foreground. The result of the thresholding is a bi-level image, where the strong edges are white and the background is black.

From this thresholded image, the centre point of all the strong edges (white pixels) are found. The assumption is that the contour of the person is symmetric, when he/she is sitting in pause-position with both hands in the lap. This means that the centre of the edge pixels describing this symmetric contour, should correspond to the centre of the figure. Around the centre line from the top of the head to the bottom of the image, a rectangle with a width of approximately 2/5 of the width of the image is defined. This region is split horizontally, giving a lower and an upper region.

To be able to identify the hands, we were here forced to make the assumption that the hands are brighter than the background when resting in the pause position. Based on this assumption, we threshold the lower region in the original colour image. The hands should then appear white, while most of the background is black. However, some of the white pixels may not be part of the hands. The white regions are therefore analysed further. Areas of connected white pixels are identified, and the size of the areas are determined. The largest area of connected white pixels are assumed to correspond to the hands. For the identified area, the circumscribed box is determined, and then the lower region is determined as a box slightly larger than this. Finally, the size of the upper region is adjusted. The next step is then to analyse the lower region to estimate a skin colour model for the hands.

Defining the colour model

When the lower region has been defined, we need to establish a model for the colour of the hands to be able to determine whether there are skin pixels in the different regions during analysis. Ideally we would like to establish a general skin colour model which could work for all videos. Colour is however not a physical phenomenon, but a perceptual phenomenon. The colour representation of skin obtained by a camera is influenced by many factors, such as ambient light, object movement etc. Different cameras produce significantly different colour values even for the same person under the same lighting condition, and human skin colours differ from person to person. General models have the advantage of being relatively user-independent, but they are not as good as user-specific models would be. In order to achieve the best segmentation results, we therefore decided to use a user-specific skin colour model.

Most video cameras use an RGB representation, however this is not necessarily the best colour representation for characterising skin-colour. The human visual system adapts to different brightness and various illumination sources such that a perception of colour is maintained within a wide range of lighting conditions. Therefore it is possible for us to remove brightness from skin-colour representation, while preserving an accurate colour information. However, in the RGB space, a triple (r,g,b) represents not only colour but also brightness. Thus, other colour models may be more suitable. There exists a lot of colour models, but we have chosen to use a simple RGB normalisation:

$$r = R/(R+G+B) \quad \text{and} \quad g = G/(R+G+B)$$

This representation defines normalised chromatic colours, and removes the brightness dependent information while preserving the colour. In addition the complexity of the RGB colour space is simplified by dimensional reduction of the colour space.

With a user-specific colour model, we will in the initialisation phase need a way to identify the pixels corresponding to the hands without having a general colour model. We have therefore assumed that the white pixels of the thresholded region of interest correspond to the hands (which is why we have made requirement number 2 previously in this section). From this a 2-dimensional histogram is computed, where the number of each colour (rg value) is counted for the hands. Later, during the analysis, this histogram can be used as a look-up-table, where the value in the histogram corresponds to the probability of a colour being a skin colour.

The definition of the regions and the colour model is part of the initialisation of the analysis of a sequence. To make it robust to changes in colour that often appears during the first few frames, not only one frame is used, but several.

Parameter estimation

In addition to the skin colour model, there are also a few other parameters that need to be estimated during the initialisation phase to adapt to the current video sequence. These are parameters defining the portion of skin pixels expected in the upper and lower region during a pause. Later, these are used to determine the thresholds needed to classify frames as break or no-break during the analysis.

The initialisation phase is necessary because there are large differences between the sequences and how the different person behave, and general estimates can not be found. The initialisation phase lasts through the first few frames, assuming the recording starts with the person sitting in the pause position. It then lasts until the person starts moving.

2.1.2 Analysis

During the analysis we use colour information to identify pixels with skin colour within each of the two identified regions of interest. The idea is that a high portion of skin coloured pixels in the lower region combined with a low portion of skin coloured pixels in the upper region, means that the hands are in the pause position. Based on the counts of skin pixels, each frame can then be classified as corresponding to a break or not.

Feature extraction

During the detection, features are extracted from each frame. The first features are based on the amount of hand pixels in each of the two regions defined during initialisation:

- The number of skin pixels in the upper region.
- The number of skin pixels in the lower region.

In order to speed up the segmentation process, a look-up table is generated which relates each normalised chromatic colour with its corresponding area inside the RGB-cube. The look-up table is used for classification of all pixels, whether they represent skin colour or not.

In addition to the number of skin pixels, a feature related to the amount of activity is extracted. This feature is however, mainly used to determine key frames (see Section 2.2):

- The difference from one frame to the next.

The analysis is based on these three features.

Classification of frames

To classify each frame based on the values of these features, there are a number of methods to choose from. A method based on hidden Markov models would probably be well suited for this problem, as this method can treat a sequence as a time series and make the classification of one frame dependent on what happened in the previous frame. Hidden Markov models have also been used for recognition of sign language by Pentland and Starner.

Due to the limitations of our study, we were however forced to use a simpler approach where we use thresholds to determine from the features whether the frame should be classified as a break or not. The thresholds are based on the parameters estimated during the initialisation, and the pre-classification is based on the amount of skin pixels in the upper and lower region. The idea is that, a high number of skin pixels in the lower region combined with a low number of pixels in the upper region, means that both hands are in the lower region and that there is a break.

When each frame has been preclassified in this way, the results are filtered to smooth the result and remove spurious errors. The breaks typically last for 12 frames and upwards. It was however, desirable to identify a break with one frame only. The middle frame of a consecutive segment of frames classified as breaks is therefore used to mark the break.

2.2 Extraction of key frames

For the extraction of key frames we have only developed a very simple method, extracting key frames based on the amount of movement. The idea is that frames with much activity are important.

Feature extraction

When finding the amount of movement in an image, the simplest approach is to use a differencing technique. This means that the difference between two consecutive frames are computed, and the amount of change between the frames can be used as a measure for movement between the two frames.

The colours of the video are not absolutely stable, and can change somewhat between frames. Some codecs will make artefacts like blocking appear in the image. This means that there will be some variation in colour or grey level between frames, even when there is no movement. Hence, rather than using the difference between two frames directly, we compute the edges in the difference image. Hence the effect of homogeneous shifts in lighting will be reduced.

The edge image is then tresholded and the foreground pixels (strong edges) are counted. Many strong edge pixels indicate much movement. We have looked only at the movement in the upper box, as much movement here means that the hands are probably in this area, and this indicates that we do not have a pause. (The opposite need not necessarily be true. There may be no movement in the upper box, without there being a pause.) The number of strong edge pixels is computed relative to the size of the upper region, and this number is then used directly as a movement feature.

Classification

When extracting the key frames, we want to extract a suitable number of frames per sequence. According to the sign language experts one sign typically takes one second. Initially they therefore suggested to extract one key frame per second. However, to get some more flexibility we instead decided to extract approximately 10 key frames per 10 seconds.

Although the movement features are computed for each frame, the selection of the key frames is not done until all the frames have been processed. The approach is as follows. For each interval of 10 seconds the amount of motion extracted as a feature from each frame, is sorted in decreasing order. Then the frames corresponding to the N highest values are selected. As several of the frames among the highest motion values will result from the same maximum, we select a number which is approximately 4 times as high as the desired number of frames for the time interval. The selected frames are marked as potential key-frames, and then a filtering is performed in the same way as for the pause frames, to select only the midpoint of a sequence of consecutive key-frames.

Chapter 3

Implementation

3.1 Platform

The program is implemented as a console application in C, using Microsoft Visual C++/C and their Video for Windows. Video for Windows is an entire system for handling video in Microsoft Windows, where the AVI file and file format is a central part.

Video for Windows also provides libraries, which can be used to manipulate frames in an AVI video stream. To be able to use the necessary AVIFile functions a link to the winmm.lib and vfw32.lib is needed. These libraries are included with MSVC++ 6.0, and should be available also for versions as low as 4.0. Newer versions of the libraries are also part of the latest Platform SDK release from Microsoft.

3.2 Syntax and parameters

The syntax of the program is as follows:

```
Usage:   signed  <AVI-file>  <Result file>
         [-Plength <pause length>] [-Pfactor <pause factor>]
         [-Nkeys <Nof key frames>] [-Nsecs <interval length>]
```

Here, the first two arguments are required as they are used to specify the input and the output file. The last four arguments are optional switches, which can be used to adjust the performance of the pause detection and key frame extraction. In the following we will describe each parameter in some more detail.

AVI-file: The name (and path) of the AVI-video to be analysed.

Result file: The name of the file which will contain the results. This will be a text file giving the identified frame types (pause or key frame), the frame number and the corresponding time code computed from the frame number. There will be one line for each classified frame, and the format of each line is as follows:

```
<Frame type>  <Frame number>  <Time-code for frame>
```

The frame type will be identified with the character P (Pause) or K (Key frame). Examples of these files can be seen in Appendix B, where we have listed the results from some experiments.

Pause length: This parameter can be used to specify the minimum length of a break. The length should be given in number of frames. If no length is specified, the parameter is set to a default value of 8 frames, which corresponds to 1/3 of a second. This seems to work well, and shorter periods than this that are classified as breaks, are often false classifications.

Pause factor: This parameter can be used to tune the number of pauses detected. The default value is 1.0. If this factor is increased, more pauses will be detected, but this also increases the possibility of detecting false pauses. If the factor is decreased fewer pauses will be detected. This can mean fewer false pauses, but also more missed pauses.

Nkeys and Nsecs: These parameters are used to specify the desired number of key frames extracted for a time interval. The default values are, Nkeys=10 and Nsecs=10, which means that there will be extracted approximately 10 key frames per 10 seconds.

Chapter 4

Experiments and Results

4.1 Data

For the experiments we had colour videos of three different persons. Each sequence was an AVI-video (Indeo 3.2) with 25 frames per second, and with a frame size of 640x480 pixels.

We had one sequence from each of the three persons. Each person was telling a different story using sign language, and the length varied for each video (see Table 4.1). The videos were all recorded with a homogeneous background. An image from each sequence can be seen in Figure 4.1. We will refer to the videos by the name of the person appearing in each video.

Name	Length	No. of frames
Georg	02:54	4361
Gudmund	03:41	5541
Torkil	11:04	16609

Table 4.1: *Lengths of the three videos used in the experiment.*

4.2 Experiments

In Figure 4.2 plots of the features extracted from the first 1500 frames of the video sequence *Georg* are shown. The upper plot shows the variation in the number of skin pixels in the lower region of interest. The dotted line in the same plot indicates the location of the true pauses (zero when no pause). From this we can see that there is a correspondence between high counts of skin pixels in the lower region and pauses. However, as can also be seen, it is difficult to determine a threshold where all frames above the threshold corresponds to a pause, and all below to action.

Torkil

Georg

Gudmund

Figure 4.1: *Pause-position for each of the persons.*

The second plot in Figure 4.2 gives the number of skin pixels in the upper region. The correspondence with the pauses is not so clear here. The count is usually low when there is a pause, but the count can also be low at other points. The third plot in Figure 4.2 shows the amount of movement between frames. Here, there is little correspondence between pauses and the levels of motion, and this is why we decided to use this feature only for the key frame extraction.

In Figure 4.3 we have shown the results of the classification for the same sequence. The upper plot shows the results of the classification. The high peaks identify the frames corresponding to the midpoints of the detected pauses, while the low peaks identify the frames where key frames were extracted. The lower plot shows the position and length of the manually identified pauses for the sequence.

The pause-classification and the key frame extraction were run on all the three video sequences. A detailed listing of the results is given in Appendix B. In Table 4.2 we have summarised the results of the pause-detection. The table gives the number of pauses that were correctly detected, and the number of falsely detected pauses. Ideally the number of correctly detected pauses should correspond to the number of manually identified pauses, while the number of falsely detected pauses should be zero.

The effects of adjusting the pause-factor is also demonstrated in Table 4.2. As expected, it can be seen that by increasing the pause factor the number of correctly classified pauses is increased, but also the number of falsely classified pauses.

	Georg			**Gudmund**			**Torkil**		
Pause factor	0.8	1.0	1.2	0.8	1.0	1.2	0.8	1.0	1.2
Manually identified pauses	**15**	**15**	**15**	**23**	**23**	**23**	**23**	**23**	**23**
Correctly detected pauses	7	10	13	21	23	23	20	22	22
Falsely detected pauses	0	1	5	1	1	4	3	7	35

Table 4.2: *Results of the pause detection for each of the three videos.*

In video analysis, speed can also be a very important factor, and for the current prototype the analysis is performed at a speed of approximately 10 frames per second on a standard NT PC.

4.3 Evaluation of results

For the break detection, the results seem quite good. From the plots of the features, we see that these features combined with hard thresholds work well but some errors are unavoidable. This means that lower thresholds, may mean that fewer pauses will be missed, but more false pauses will be detected. A higher threshold means that the risk of detecting false pauses decreases, but then more pauses may be missed.

The results could probably be further improved by using a statistical classification method. The hidden Markov model discussed earlier could for instance be used to incorporate information from more than one frame. It would be very interesting to study these techniques more closely for this application. The initial results obtained with our very simple classification schemes, indicate that even better results could be obtained with more sophisticated methods.

For the key frame extraction, we need sign language expertise to evaluate the results, but the initial feedback on the results indicate that they seem very interesting. However, closer evaluation is necessary.

The test videos in these experiments should represent the expected quality of home made sign language videos. The three sequences we have used are also quite different in how the persons use their hands. Still, it should be noted that the methods have been trained and tested on the same set, and experiments on larger sets should be performed.

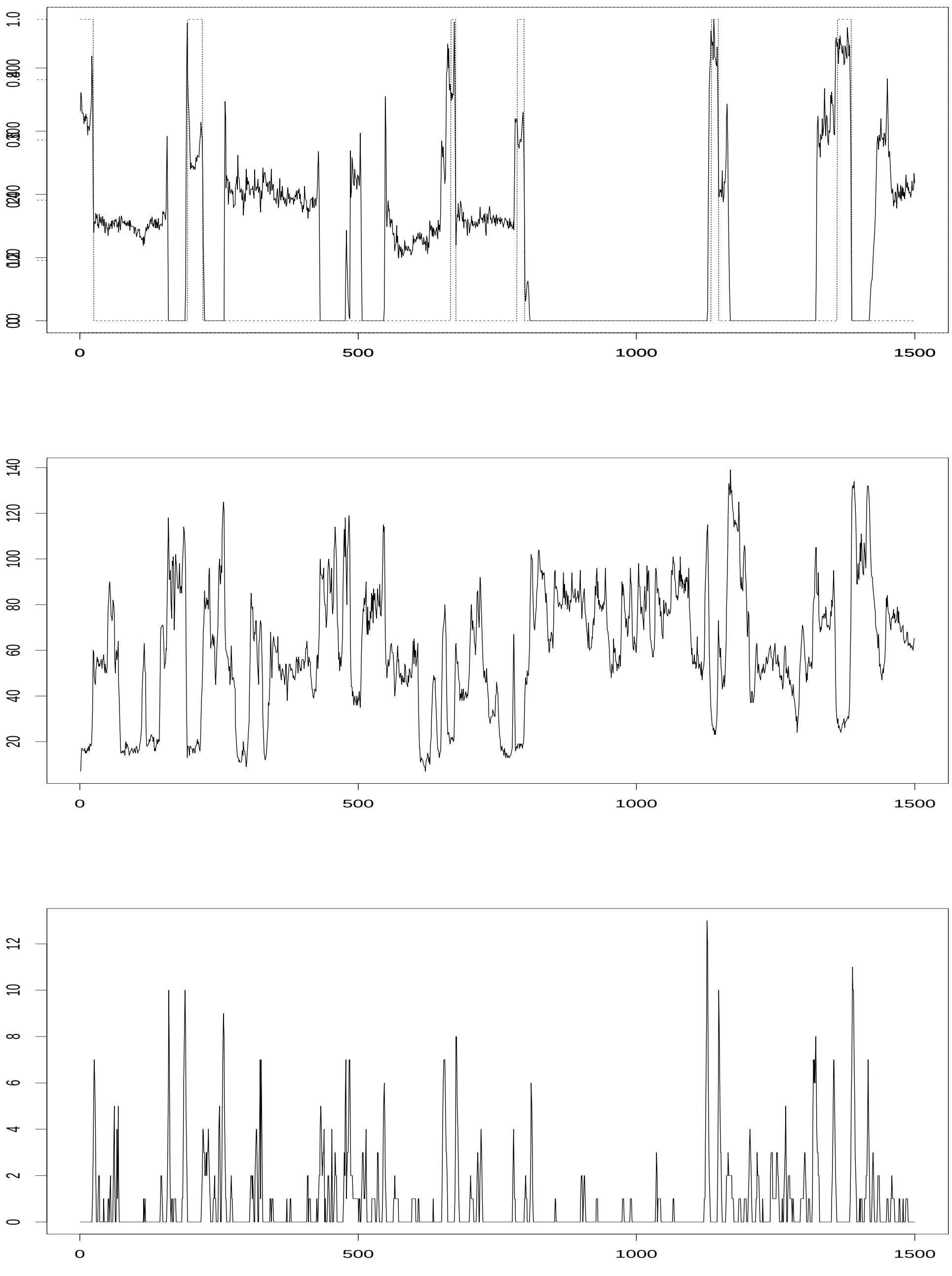


Figure 4.2: **Top:** *Number of skin pixels in the lower region and the pauses (dotted line),* **Middle:** *Number of skin-pixels in upper region,* **Bottom:** *Movement.*

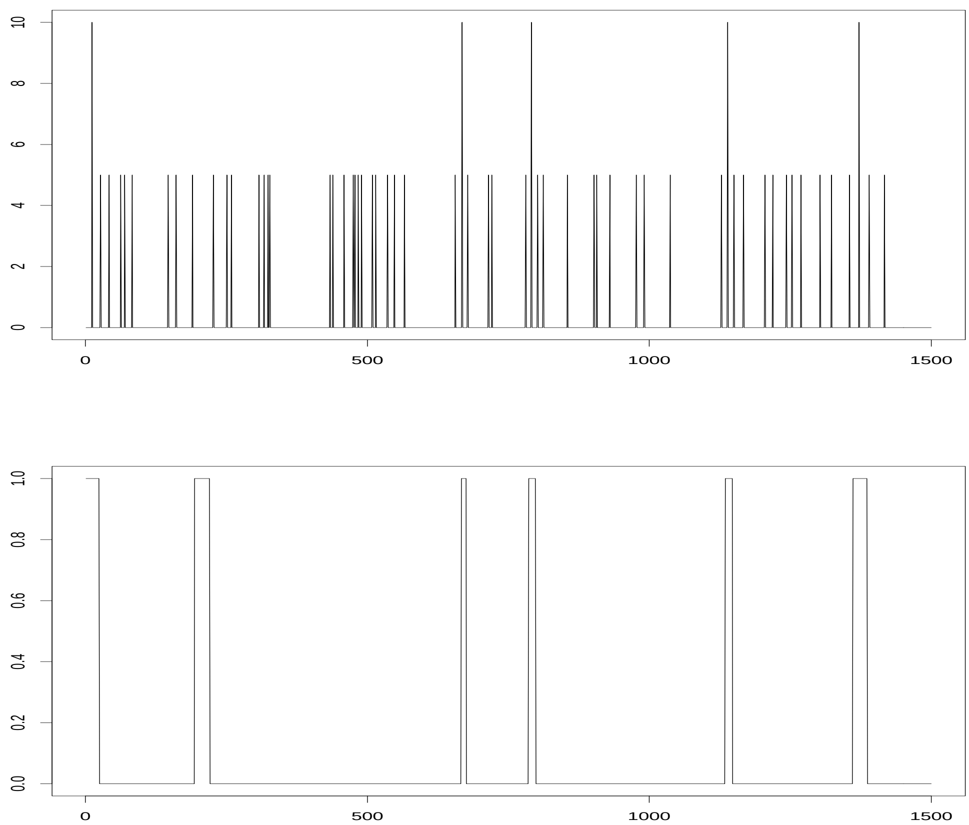


Figure 4.3: **Top:** *Identified pauses and extracted keyframes.* **Bottom:** *Manually identified pauses.*

Chapter 5

Summary and Discussion

In this report we have described an initial study on the analysis of sign language videos based on image analysis. The purpose of this study has been to perform a preliminary study of potential approaches to the problem of detecting breaks and extracting highlights from sign language videos.

The reason for looking into this problem is that Møller Kompetansesenter in Trondheim, which is a centre that develops educational tools and materials for deaf people, wants to develop a sign language processor for the deaf with a flexibility and functionality comparable to that of word processors. In addition to basic video editing tools, they want such a system to offer higher level tools that can produce an overview of the contents by automatically extracting breaks and highlights.

This has been a very limited study, and our aim has been to find out through some initial experiments whether this is a problem that can be solved using appropriate methods and to implement a simple prototype. There are two problems to be solved. First we want to detect breaks to split the video up in natural shorter sequences. Then we want to extract key frames that can help to give an overview of the contents of the sequence. In this study we have focused on the first part of the problem, i.e. to detect breaks. However, some initial experiments have also been performed with the second problem

An important characteristic of a break is that when using sign language, the person will rest with his/her hands in the lap to mark a pause. The detection of breaks is therefore based on detecting the frames where the hands are at rest in the lap. To accomplish this we have used a simple approach where we, rather than identifying the exact position of the hands, identify a few regions of interest and decide whether the hands are present in these regions. To be able to do this we made the following requirements: (i) the background is homogeneous and stationary, (ii) the person is sitting in the pause position when the video starts, and (iii) the background is darker than the hands when resting in the lap of the person.

During an initialisation phase, two regions of interest are determined and a user-specific skin-colour model is estimated. The first region of interest is chosen as the area covered by the hands during a pause, while the other region is the area where the hands move during signing. During the analysis the colour model is used to identify pixels with skin colour within each of the identified regions of interest. The idea is that a high portion of skin coloured pixels in the pause region combined with a low portion of skin coloured pixels in the other region, means that the hands are in the pause position. Based on the counts of skin pixels, each frame can then be classified as corresponding to a break or not. For the extraction of key frames we have only developed a very simple method, extracting key frames based on the amount of movement. The idea is that frames with much activity are important.

Some experiments on three video sequences with three different people, have been performed. For the break detection, the results seem quite good. The classification is however based on hard thresholds and some errors are therefore unavoidable. Even better results could possibly be obtained using a statistical classification method. For the key frame extraction, we need sign language expertise to evaluate the results, but initial feedback on the results indicate that they seem interesting.

The test videos in these experiments should represent the expected quality of home made sign language videos, and the three sequences we have used are quite different in how the persons use their hands. Still, it should be noted that the methods have been trained and tested on the same set, and experiments on larger sets should be performed. Future work should also include a study of more sophisticated techniques for motion tracking and for classification. The good results obtained with simple approaches, indicate that break detection should be possible. When it comes to key frame extraction a closer evaluation of the results is however needed before we can draw a conclusion.

Appendix A

Software

The software of this prototype consists of the following files:

C-files:

- AVIparse.c
- AVItools.c
- RGBmethods.c
- RGBtools.c
- colour.c
- region.c
- sobel.c
- readswitch.c
- mmatrix.c

h-files:

- GENdef.h
- AVIdef.h
- RGBdef.h
- region.h

In addition Video for Windows is required.

Appendix B

Results

In the following we will give some results from the experiments. These are listings of the result files for each of the three videos **Georg**, **Gudmund**, and **Torkil**.

The results were obtained with the default setting of parameters:

Plength	=	8
Pfactor	=	1.0
Nsecs	=	10
Nkeys	=	10

The results contain one line for each detected pause and each key frame. The frame type is identified with the character P (Pause) or K (Key frame), and then the frame number and the corresponding time code are given for the identified frames.

Georg

P	13	0:00:00:52
K	28	0:00:01:12
K	43	0:00:01:72
K	64	0:00:02:56
K	71	0:00:02:84
K	84	0:00:03:36
K	148	0:00:05:92
K	162	0:00:06:48
K	191	0:00:07:64
K	228	0:00:09:12
K	252	0:00:10:08
K	260	0:00:10:40
K	309	0:00:12:36
K	318	0:00:12:72
K	325	0:00:13:00
K	328	0:00:13:12
K	435	0:00:17:40
K	440	0:00:17:60
K	460	0:00:18:40
K	476	0:00:19:04
K	479	0:00:19:16
K	485	0:00:19:40
K	491	0:00:19:64
K	510	0:00:20:40
K	516	0:00:20:64
K	537	0:00:21:48
K	549	0:00:21:96
K	567	0:00:22:68
K	657	0:00:26:28
P	669	0:00:26:76
K	679	0:00:27:16
K	716	0:00:28:64
K	722	0:00:28:88
K	782	0:00:31:28
P	792	0:00:31:68
K	803	0:00:32:12
K	813	0:00:32:52
K	856	0:00:34:24
K	903	0:00:36:12
K	908	0:00:36:32
K	931	0:00:37:24
K	978	0:00:39:12
K	992	0:00:39:68
K	1038	0:00:41:52
K	1129	0:00:45:16
P	1140	0:00:45:60
K	1151	0:00:46:04
K	1168	0:00:46:72
K	1206	0:00:48:24
K	1220	0:00:48:80
K	1244	0:00:49:76
K	1254	0:00:50:16
K	1270	0:00:50:80
K	1304	0:00:52:16
K	1324	0:00:52:96
K	1356	0:00:54:24
P	1373	0:00:54:92
K	1391	0:00:55:64
K	1418	0:00:56:72
K	1517	0:01:00:68
K	1525	0:01:01:00
K	1536	0:01:01:44
K	1547	0:01:01:88
K	1553	0:01:02:12
P	1588	0:01:03:52
K	1619	0:01:04:76
K	1632	0:01:05:28
K	1669	0:01:06:76
K	1722	0:01:08:88
K	1740	0:01:09:60
K	1775	0:01:11:00
K	1808	0:01:12:32
K	1823	0:01:12:92
K	1836	0:01:13:44
K	1876	0:01:15:04
K	1905	0:01:16:20
K	1937	0:01:17:48
K	1949	0:01:17:96
P	1964	0:01:18:56
K	1980	0:01:19:20
K	1999	0:01:19:96
K	2039	0:01:21:56
P	2073	0:01:22:92
K	2106	0:01:24:24
K	2115	0:01:24:60
K	2139	0:01:25:56
K	2146	0:01:25:84
K	2149	0:01:25:96
K	2185	0:01:27:40
K	2188	0:01:27:52
K	2237	0:01:29:48
K	2245	0:01:29:80
K	2257	0:01:30:28
K	2318	0:01:32:72
K	2335	0:01:33:40
K	2375	0:01:35:00
K	2380	0:01:35:20
K	2422	0:01:36:88

K	2439	0:01:37:56
K	2452	0:01:38:08
K	2469	0:01:38:76
K	2582	0:01:43:28
K	2593	0:01:43:72
K	2624	0:01:44:96
K	2636	0:01:45:44
K	2687	0:01:47:48
K	2695	0:01:47:80
K	2723	0:01:48:92
K	2740	0:01:49:60
K	2749	0:01:49:96
K	2796	0:01:51:84
K	2832	0:01:53:28
K	2837	0:01:53:48
K	2841	0:01:53:64
P	2854	0:01:54:16
P	2885	0:01:55:40
K	2925	0:01:57:00
K	2941	0:01:57:64
K	2949	0:01:57:96
K	2959	0:01:58:36
K	2969	0:01:58:76
K	2997	0:01:59:88
K	3022	0:02:00:88
K	3028	0:02:01:12
K	3035	0:02:01:40
K	3076	0:02:03:04
K	3082	0:02:03:28
K	3088	0:02:03:52
K	3096	0:02:03:84
K	3110	0:02:04:40
K	3131	0:02:05:24
K	3159	0:02:06:36
K	3273	0:02:10:92
K	3279	0:02:11:16
K	3282	0:02:11:28
K	3289	0:02:11:56
K	3296	0:02:11:84
K	3304	0:02:12:16
K	3341	0:02:13:64
K	3360	0:02:14:40
P	3388	0:02:15:52
K	3415	0:02:16:60
K	3486	0:02:19:44
K	3497	0:02:19:88
K	3540	0:02:21:60
K	3555	0:02:22:20
K	3593	0:02:23:72
K	3608	0:02:24:32
K	3632	0:02:25:28
K	3659	0:02:26:36
K	3683	0:02:27:32
K	3703	0:02:28:12
K	3731	0:02:29:24
K	3752	0:02:30:08
P	3765	0:02:30:60
K	3779	0:02:31:16
K	3811	0:02:32:44
K	3852	0:02:34:08
K	3870	0:02:34:80
K	3885	0:02:35:40
K	3897	0:02:35:88
K	3929	0:02:37:16
K	3961	0:02:38:44
K	4039	0:02:41:56
K	4120	0:02:44:80
K	4139	0:02:45:56
K	4152	0:02:46:08
K	4171	0:02:46:84
K	4176	0:02:47:04
K	4183	0:02:47:32
K	4190	0:02:47:60
K	4196	0:02:47:84
K	4202	0:02:48:08
K	4231	0:02:49:24
K	4241	0:02:49:64

Gudmund

P	8	0:00:00:32
K	19	0:00:00:76
K	31	0:00:01:24
K	36	0:00:01:44
K	105	0:00:04:20
K	112	0:00:04:48
K	125	0:00:05:00
P	142	0:00:05:68
K	157	0:00:06:28
K	185	0:00:07:40
K	230	0:00:09:20
K	252	0:00:10:08
K	269	0:00:10:76
K	279	0:00:11:16
K	319	0:00:12:76
P	334	0:00:13:36
K	347	0:00:13:88
K	415	0:00:16:60
K	482	0:00:19:28
K	520	0:00:20:80
K	528	0:00:21:12
K	532	0:00:21:28
K	564	0:00:22:56
K	614	0:00:24:56
P	624	0:00:24:96
K	640	0:00:25:60
K	671	0:00:26:84
K	680	0:00:27:20
K	689	0:00:27:56
K	715	0:00:28:60
K	728	0:00:29:12
K	741	0:00:29:64
K	752	0:00:30:08
P	758	0:00:30:32
K	822	0:00:32:88
K	829	0:00:33:16
K	839	0:00:33:56
K	846	0:00:33:84
K	851	0:00:34:04
K	879	0:00:35:16
K	884	0:00:35:36
P	990	0:00:39:60
K	1099	0:00:43:96
K	1106	0:00:44:24
K	1111	0:00:44:44
K	1114	0:00:44:56
K	1136	0:00:45:44
K	1180	0:00:47:20
K	1184	0:00:47:36
K	1188	0:00:47:52
K	1201	0:00:48:04
P	1210	0:00:48:40
P	1227	0:00:49:08
K	1242	0:00:49:68
K	1269	0:00:50:76
K	1303	0:00:52:12
K	1320	0:00:52:80
K	1336	0:00:53:44
K	1346	0:00:53:84
K	1385	0:00:55:40
P	1400	0:00:56:00
K	1414	0:00:56:56
K	1560	0:01:02:40
K	1599	0:01:03:96
K	1606	0:01:04:24
K	1633	0:01:05:32
K	1656	0:01:06:24
K	1685	0:01:07:40
K	1691	0:01:07:64
K	1706	0:01:08:24
P	1724	0:01:08:96
K	1741	0:01:09:64
K	1756	0:01:10:24
K	1786	0:01:11:44
K	1802	0:01:12:08
K	1812	0:01:12:48
K	1816	0:01:12:64
K	1833	0:01:13:32
K	1889	0:01:15:56
P	1904	0:01:16:16
K	1918	0:01:16:72
K	1948	0:01:17:92
K	1955	0:01:18:20
K	1980	0:01:19:20
K	2029	0:01:21:16
K	2060	0:01:22:40
K	2065	0:01:22:60
K	2079	0:01:23:16
K	2138	0:01:25:52
K	2158	0:01:26:32
K	2194	0:01:27:76
K	2228	0:01:29:12
K	2235	0:01:29:40
K	2259	0:01:30:36
P	2273	0:01:30:92
K	2290	0:01:31:60
K	2348	0:01:33:92
K	2384	0:01:35:36

K	2407	0:01:36:28
K	2422	0:01:36:88
P	2461	0:01:38:44
K	2470	0:01:38:80
K	2478	0:01:39:12
K	2484	0:01:39:36
K	2505	0:01:40:20
K	2524	0:01:40:96
K	2544	0:01:41:76
K	2553	0:01:42:12
K	2576	0:01:43:04
K	2588	0:01:43:52
K	2606	0:01:44:24
K	2632	0:01:45:28
K	2660	0:01:46:40
P	2693	0:01:47:72
K	2725	0:01:49:00
K	2761	0:01:50:44
P	2773	0:01:50:92
K	2789	0:01:51:56
K	2853	0:01:54:12
K	2863	0:01:54:52
K	2870	0:01:54:80
K	2894	0:01:55:76
K	2904	0:01:56:16
K	2917	0:01:56:68
K	2957	0:01:58:28
K	3047	0:02:01:88
K	3052	0:02:02:08
K	3084	0:02:03:36
K	3091	0:02:03:64
K	3095	0:02:03:80
K	3118	0:02:04:72
K	3140	0:02:05:60
P	3164	0:02:06:56
K	3187	0:02:07:48
K	3218	0:02:08:72
K	3235	0:02:09:40
K	3240	0:02:09:60
P	3278	0:02:11:12
K	3307	0:02:12:28
K	3333	0:02:13:32
K	3349	0:02:13:96
K	3408	0:02:16:32
K	3417	0:02:16:68
K	3427	0:02:17:08
K	3469	0:02:18:76
P	3483	0:02:19:32
K	3498	0:02:19:92
K	3509	0:02:20:36
K	3546	0:02:21:84
K	3561	0:02:22:44
K	3584	0:02:23:36
K	3620	0:02:24:80
P	3632	0:02:25:28
K	3642	0:02:25:68
K	3686	0:02:27:44
K	3692	0:02:27:68
K	3709	0:02:28:36
K	3759	0:02:30:36
K	3786	0:02:31:44
K	3806	0:02:32:24
K	3843	0:02:33:72
K	3862	0:02:34:48
P	3880	0:02:35:20
K	3890	0:02:35:60
K	3908	0:02:36:32
K	3925	0:02:37:00
K	3947	0:02:37:88
K	3958	0:02:38:32
K	3970	0:02:38:80
K	3975	0:02:39:00
K	3997	0:02:39:88
K	4001	0:02:40:04
K	4005	0:02:40:20
K	4012	0:02:40:48
K	4043	0:02:41:72
K	4054	0:02:42:16
K	4157	0:02:46:28
K	4216	0:02:48:64
K	4239	0:02:49:56
K	4257	0:02:50:28
K	4278	0:02:51:12
P	4290	0:02:51:60
K	4299	0:02:51:96
P	4326	0:02:53:04
K	4347	0:02:53:88
K	4356	0:02:54:24
K	4378	0:02:55:12
K	4434	0:02:57:36
K	4466	0:02:58:64
P	4493	0:02:59:72
K	4517	0:03:00:68
K	4553	0:03:02:12
K	4568	0:03:02:72
K	4621	0:03:04:84
K	4641	0:03:05:64
K	4647	0:03:05:88
K	4653	0:03:06:12

K	4696	0:03:07:84
K	4734	0:03:09:36
K	4747	0:03:09:88
K	4754	0:03:10:16
K	4803	0:03:12:12
K	4808	0:03:12:32
K	4819	0:03:12:76
K	4826	0:03:13:04
K	4837	0:03:13:48
K	4861	0:03:14:44
P	4888	0:03:15:52
P	4943	0:03:17:72
K	4952	0:03:18:08
K	4963	0:03:18:52
K	4989	0:03:19:56
K	4997	0:03:19:88
K	5020	0:03:20:80
K	5039	0:03:21:56
K	5045	0:03:21:80
K	5093	0:03:23:72
K	5135	0:03:25:40
K	5149	0:03:25:96
K	5156	0:03:26:24
K	5166	0:03:26:64
P	5197	0:03:27:88
K	5228	0:03:29:12
K	5244	0:03:29:76
K	5252	0:03:30:08
K	5285	0:03:31:40
K	5312	0:03:32:48
K	5339	0:03:33:56
K	5380	0:03:35:20
P	5398	0:03:35:92
K	5415	0:03:36:60
K	5425	0:03:37:00
K	5444	0:03:37:76
K	5486	0:03:39:44

Torkil

P	3	0:00:00:12
K	9	0:00:00:36
K	14	0:00:00:56
K	22	0:00:00:88
K	26	0:00:01:04
K	36	0:00:01:44
K	40	0:00:01:60
K	50	0:00:02:00
K	81	0:00:03:24
K	102	0:00:04:08
K	179	0:00:07:16
K	185	0:00:07:40
K	199	0:00:07:96
K	237	0:00:09:48
K	372	0:00:14:88
K	378	0:00:15:12
K	383	0:00:15:32
K	396	0:00:15:84
K	429	0:00:17:16
P	465	0:00:18:60
K	499	0:00:19:96
K	511	0:00:20:44
P	524	0:00:20:96
K	537	0:00:21:48
K	566	0:00:22:64
K	624	0:00:24:96
K	679	0:00:27:16
K	684	0:00:27:36
K	690	0:00:27:60
K	694	0:00:27:76
K	697	0:00:27:88
K	718	0:00:28:72
K	724	0:00:28:96
K	738	0:00:29:52
K	745	0:00:29:80
P	761	0:00:30:44
K	779	0:00:31:16
K	873	0:00:34:92
K	886	0:00:35:44
K	896	0:00:35:84
P	910	0:00:36:40
K	924	0:00:36:96
K	933	0:00:37:32
K	992	0:00:39:68
K	998	0:00:39:92
K	1037	0:00:41:48
K	1049	0:00:41:96
K	1053	0:00:42:12

K	1061	0:00:42:44
K	1066	0:00:42:64
K	1070	0:00:42:80
P	1083	0:00:43:32
K	1096	0:00:43:84
K	1169	0:00:46:76
K	1217	0:00:48:68
K	1244	0:00:49:76
K	1252	0:00:50:08
K	1293	0:00:51:72
K	1299	0:00:51:96
K	1338	0:00:53:52
P	1352	0:00:54:08
K	1365	0:00:54:60
K	1386	0:00:55:44
P	1401	0:00:56:04
K	1410	0:00:56:40
K	1480	0:00:59:20
K	1505	0:01:00:20
K	1570	0:01:02:80
K	1577	0:01:03:08
K	1582	0:01:03:28
K	1602	0:01:04:08
K	1621	0:01:04:84
K	1631	0:01:05:24
K	1635	0:01:05:40
K	1641	0:01:05:64
P	1660	0:01:06:40
P	1681	0:01:07:24
K	1699	0:01:07:96
K	1756	0:01:10:24
K	1771	0:01:10:84
K	1792	0:01:11:68
K	1842	0:01:13:68
K	1863	0:01:14:52
K	1896	0:01:15:84
K	1907	0:01:16:28
K	1916	0:01:16:64
K	1941	0:01:17:64
K	1947	0:01:17:88
K	1952	0:01:18:08
K	1968	0:01:18:72
K	1984	0:01:19:36
K	2022	0:01:20:88
K	2032	0:01:21:28
K	2047	0:01:21:88
K	2062	0:01:22:48
K	2104	0:01:24:16
K	2112	0:01:24:48
K	2139	0:01:25:56
P	2153	0:01:26:12
K	2182	0:01:27:28
K	2202	0:01:28:08
K	2213	0:01:28:52
K	2241	0:01:29:64
K	2291	0:01:31:64
K	2300	0:01:32:00
K	2336	0:01:33:44
K	2366	0:01:34:64
K	2403	0:01:36:12
K	2409	0:01:36:36
K	2415	0:01:36:60
K	2441	0:01:37:64
K	2449	0:01:37:96
K	2459	0:01:38:36
K	2462	0:01:38:48
K	2473	0:01:38:92
K	2491	0:01:39:64
K	2532	0:01:41:28
K	2555	0:01:42:20
K	2598	0:01:43:92
K	2603	0:01:44:12
K	2737	0:01:49:48
K	2745	0:01:49:80
K	2750	0:01:50:00
K	2765	0:01:50:60
K	2778	0:01:51:12
K	2783	0:01:51:32
K	2803	0:01:52:12
K	2823	0:01:52:92
K	2834	0:01:53:36
K	2850	0:01:54:00
K	2932	0:01:57:28
K	2946	0:01:57:84
K	3095	0:02:03:80
K	3125	0:02:05:00
K	3173	0:02:06:92
K	3188	0:02:07:52
K	3199	0:02:07:96
K	3204	0:02:08:16
K	3221	0:02:08:84
K	3228	0:02:09:12
K	3231	0:02:09:24
K	3234	0:02:09:36
K	3279	0:02:11:16
K	3293	0:02:11:72
K	3302	0:02:12:08
K	3326	0:02:13:04

K	3331	0:02:13:24
K	3363	0:02:14:52
K	3368	0:02:14:72
K	3420	0:02:16:80
K	3425	0:02:17:00
K	3441	0:02:17:64
K	3476	0:02:19:04
K	3504	0:02:20:16
K	3513	0:02:20:52
K	3528	0:02:21:12
K	3537	0:02:21:48
K	3562	0:02:22:48
K	3569	0:02:22:76
K	3584	0:02:23:36
K	3671	0:02:26:84
K	3679	0:02:27:16
K	3684	0:02:27:36
K	3781	0:02:31:24
K	3807	0:02:32:28
K	3864	0:02:34:56
K	3886	0:02:35:44
K	3907	0:02:36:28
K	3921	0:02:36:84
K	3931	0:02:37:24
K	3946	0:02:37:84
K	4022	0:02:40:88
K	4027	0:02:41:08
K	4042	0:02:41:68
K	4060	0:02:42:40
K	4104	0:02:44:16
K	4183	0:02:47:32
K	4240	0:02:49:60
P	4247	0:02:49:88
K	4264	0:02:50:56
K	4326	0:02:53:04
K	4330	0:02:53:20
K	4366	0:02:54:64
K	4373	0:02:54:92
K	4391	0:02:55:64
K	4399	0:02:55:96
K	4409	0:02:56:36
K	4420	0:02:56:80
K	4476	0:02:59:04
P	4500	0:03:00:00
K	4519	0:03:00:76
K	4531	0:03:01:24
K	4562	0:03:02:48
K	4629	0:03:05:16
K	4650	0:03:06:00
K	4680	0:03:07:20
K	4693	0:03:07:72
P	4703	0:03:08:12
K	4712	0:03:08:48
K	4738	0:03:09:52
K	4744	0:03:09:76
K	4751	0:03:10:04
K	4768	0:03:10:72
K	4771	0:03:10:84
K	4776	0:03:11:04
K	4786	0:03:11:44
P	4813	0:03:12:52
K	4843	0:03:13:72
K	4848	0:03:13:92
K	4862	0:03:14:48
K	4909	0:03:16:36
K	4928	0:03:17:12
K	4973	0:03:18:92
K	5010	0:03:20:40
K	5018	0:03:20:72
K	5148	0:03:25:92
K	5153	0:03:26:12
K	5209	0:03:28:36
K	5219	0:03:28:76
K	5222	0:03:28:88
K	5227	0:03:29:08
K	5248	0:03:29:92
K	5266	0:03:30:64
K	5277	0:03:31:08
K	5296	0:03:31:84
K	5300	0:03:32:00
K	5323	0:03:32:92
K	5340	0:03:33:60
K	5366	0:03:34:64
K	5379	0:03:35:16
K	5388	0:03:35:52
K	5455	0:03:38:20
K	5522	0:03:40:88
K	5526	0:03:41:04
K	5551	0:03:42:04
K	5568	0:03:42:72
K	5587	0:03:43:48
K	5616	0:03:44:64
K	5652	0:03:46:08
K	5667	0:03:46:68
K	5731	0:03:49:24
K	5767	0:03:50:68
K	5779	0:03:51:16
K	5815	0:03:52:60
K	5819	0:03:52:76

K	5908	0:03:56:32
K	5932	0:03:57:28
K	5946	0:03:57:84
K	5970	0:03:58:80
K	5982	0:03:59:28
K	6020	0:04:00:80
K	6026	0:04:01:04
K	6054	0:04:02:16
K	6062	0:04:02:48
K	6116	0:04:04:64
K	6138	0:04:05:52
K	6164	0:04:06:56
K	6171	0:04:06:84
K	6177	0:04:07:08
K	6181	0:04:07:24
K	6211	0:04:08:44
K	6288	0:04:11:52
K	6325	0:04:13:00
K	6332	0:04:13:28
K	6370	0:04:14:80
K	6386	0:04:15:44
K	6392	0:04:15:68
K	6424	0:04:16:96
K	6451	0:04:18:04
K	6480	0:04:19:20
K	6489	0:04:19:56
K	6538	0:04:21:52
K	6589	0:04:23:56
K	6598	0:04:23:92
K	6603	0:04:24:12
K	6621	0:04:24:84
K	6647	0:04:25:88
K	6671	0:04:26:84
K	6680	0:04:27:20
K	6693	0:04:27:72
K	6697	0:04:27:88
K	6726	0:04:29:04
K	6756	0:04:30:24
K	6770	0:04:30:80
K	6811	0:04:32:44
K	6921	0:04:36:84
K	6937	0:04:37:48
K	6951	0:04:38:04
K	6959	0:04:38:36
K	6967	0:04:38:68
K	6970	0:04:38:80
K	6975	0:04:39:00
K	6989	0:04:39:56
K	6997	0:04:39:88
K	7047	0:04:41:88
K	7058	0:04:42:32
K	7112	0:04:44:48
K	7166	0:04:46:64
K	7170	0:04:46:80
K	7180	0:04:47:20
K	7185	0:04:47:40
K	7190	0:04:47:60
K	7208	0:04:48:32
K	7214	0:04:48:56
K	7223	0:04:48:92
K	7239	0:04:49:56
K	7249	0:04:49:96
K	7266	0:04:50:64
K	7276	0:04:51:04
K	7287	0:04:51:48
K	7334	0:04:53:36
K	7352	0:04:54:08
K	7382	0:04:55:28
K	7436	0:04:57:44
K	7441	0:04:57:64
K	7446	0:04:57:84
K	7451	0:04:58:04
K	7467	0:04:58:68
K	7512	0:05:00:48
K	7521	0:05:00:84
K	7538	0:05:01:52
K	7547	0:05:01:88
K	7553	0:05:02:12
K	7614	0:05:04:56
K	7630	0:05:05:20
K	7647	0:05:05:88
K	7673	0:05:06:92
K	7704	0:05:08:16
K	7777	0:05:11:08
K	7787	0:05:11:48
K	7798	0:05:11:92
K	7803	0:05:12:12
K	7811	0:05:12:44
K	7850	0:05:14:00
K	7889	0:05:15:56
K	7921	0:05:16:84
K	7926	0:05:17:04
K	7930	0:05:17:20
K	7954	0:05:18:16
K	7961	0:05:18:44
K	7965	0:05:18:60
K	7969	0:05:18:76
K	8007	0:05:20:28
K	8013	0:05:20:52

K	8025	0:05:21:00
K	8047	0:05:21:88
K	8052	0:05:22:08
K	8075	0:05:23:00
K	8155	0:05:26:20
K	8165	0:05:26:60
K	8225	0:05:29:00
K	8234	0:05:29:36
K	8278	0:05:31:12
P	8315	0:05:32:60
K	8349	0:05:33:96
K	8356	0:05:34:24
K	8361	0:05:34:44
K	8389	0:05:35:56
K	8399	0:05:35:96
K	8405	0:05:36:20
K	8408	0:05:36:32
K	8447	0:05:37:88
K	8494	0:05:39:76
K	8523	0:05:40:92
K	8541	0:05:41:64
K	8591	0:05:43:64
K	8628	0:05:45:12
K	8638	0:05:45:52
K	8656	0:05:46:24
K	8683	0:05:47:32
K	8718	0:05:48:72
K	8729	0:05:49:16
K	8755	0:05:50:20
K	8760	0:05:50:40
K	8774	0:05:50:96
K	8790	0:05:51:60
K	8822	0:05:52:88
K	8888	0:05:55:52
K	8935	0:05:57:40
K	8985	0:05:59:40
K	8991	0:05:59:64
K	8999	0:05:59:96
K	9021	0:06:00:84
K	9028	0:06:01:12
K	9039	0:06:01:56
K	9061	0:06:02:44
K	9079	0:06:03:16
K	9112	0:06:04:48
P	9139	0:06:05:56
K	9165	0:06:06:60
K	9178	0:06:07:12
K	9250	0:06:10:00
K	9257	0:06:10:28
K	9343	0:06:13:72
K	9354	0:06:14:16
K	9369	0:06:14:76
K	9398	0:06:15:92
K	9406	0:06:16:24
K	9461	0:06:18:44
K	9477	0:06:19:08
K	9483	0:06:19:32
K	9501	0:06:20:04
K	9507	0:06:20:28
K	9527	0:06:21:08
K	9587	0:06:23:48
K	9605	0:06:24:20
K	9623	0:06:24:92
K	9632	0:06:25:28
K	9743	0:06:29:72
K	9757	0:06:30:28
K	9766	0:06:30:64
K	9772	0:06:30:88
K	9803	0:06:32:12
K	9818	0:06:32:72
K	9877	0:06:35:08
P	9886	0:06:35:44
K	9894	0:06:35:76
K	9929	0:06:37:16
P	9955	0:06:38:20
K	9978	0:06:39:12
K	10079	0:06:43:16
P	10090	0:06:43:60
K	10099	0:06:43:96
K	10121	0:06:44:84
K	10126	0:06:45:04
K	10134	0:06:45:36
K	10137	0:06:45:48
K	10170	0:06:46:80
K	10175	0:06:47:00
K	10184	0:06:47:36
K	10210	0:06:48:40
K	10230	0:06:49:20
K	10246	0:06:49:84
K	10249	0:06:49:96
K	10281	0:06:51:24
K	10321	0:06:52:84
P	10349	0:06:53:96
K	10375	0:06:55:00
K	10407	0:06:56:28
K	10417	0:06:56:68
K	10436	0:06:57:44
K	10464	0:06:58:56
K	10495	0:06:59:80
K	10523	0:07:00:92

K	10534	0:07:01:36
K	10542	0:07:01:68
K	10548	0:07:01:92
K	10552	0:07:02:08
K	10579	0:07:03:16
K	10661	0:07:06:44
P	10688	0:07:07:52
K	10713	0:07:08:52
K	10776	0:07:11:04
P	10784	0:07:11:36
K	10791	0:07:11:64
K	10800	0:07:12:00
K	10809	0:07:12:36
K	10818	0:07:12:72
K	10838	0:07:13:52
P	10877	0:07:15:08
K	10952	0:07:18:08
K	10985	0:07:19:40
K	10995	0:07:19:80
K	11007	0:07:20:28
P	11017	0:07:20:68
K	11027	0:07:21:08
K	11045	0:07:21:80
K	11062	0:07:22:48
K	11096	0:07:23:84
K	11123	0:07:24:92
K	11187	0:07:27:48
K	11196	0:07:27:84
K	11206	0:07:28:24
K	11212	0:07:28:48
K	11260	0:07:30:40
K	11269	0:07:30:76
K	11315	0:07:32:60
K	11413	0:07:36:52
P	11435	0:07:37:40
K	11457	0:07:38:28
K	11476	0:07:39:04
K	11487	0:07:39:48
K	11491	0:07:39:64
K	11494	0:07:39:76
K	11498	0:07:39:92
K	11509	0:07:40:36
K	11517	0:07:40:68
K	11521	0:07:40:84
K	11541	0:07:41:64
K	11555	0:07:42:20
K	11566	0:07:42:64
K	11679	0:07:47:16
K	11701	0:07:48:04
K	11714	0:07:48:56
K	11751	0:07:50:04
K	11815	0:07:52:60
K	11835	0:07:53:40
K	11852	0:07:54:08
K	11856	0:07:54:24
K	11900	0:07:56:00
K	11904	0:07:56:16
K	11907	0:07:56:28
K	11910	0:07:56:40
K	11913	0:07:56:52
K	11918	0:07:56:72
K	11931	0:07:57:24
K	11992	0:07:59:68
K	12006	0:08:00:24
K	12048	0:08:01:92
P	12085	0:08:03:40
K	12119	0:08:04:76
K	12133	0:08:05:32
K	12137	0:08:05:48
K	12173	0:08:06:92
K	12178	0:08:07:12
K	12183	0:08:07:32
K	12190	0:08:07:60
K	12225	0:08:09:00
K	12236	0:08:09:44
K	12304	0:08:12:16
K	12314	0:08:12:56
K	12320	0:08:12:80
K	12327	0:08:13:08
K	12337	0:08:13:48
K	12342	0:08:13:68
K	12371	0:08:14:84
K	12376	0:08:15:04
K	12423	0:08:16:92
K	12477	0:08:19:08
K	12484	0:08:19:36
K	12499	0:08:19:96
K	12537	0:08:21:48
K	12544	0:08:21:76
K	12562	0:08:22:48
K	12566	0:08:22:64
K	12683	0:08:27:32
K	12719	0:08:28:76
K	12734	0:08:29:36
K	12837	0:08:33:48
K	12882	0:08:35:28
K	12895	0:08:35:80
K	12899	0:08:35:96
K	12909	0:08:36:36
K	12923	0:08:36:92

K	12926	0:08:37:04
K	12934	0:08:37:36
K	12946	0:08:37:84
K	12976	0:08:39:04
K	12998	0:08:39:92
K	13030	0:08:41:20
K	13034	0:08:41:36
K	13147	0:08:45:88
K	13196	0:08:47:84
K	13201	0:08:48:04
K	13206	0:08:48:24
K	13220	0:08:48:80
K	13232	0:08:49:28
K	13243	0:08:49:72
K	13250	0:08:50:00
K	13265	0:08:50:60
K	13279	0:08:51:16
K	13311	0:08:52:44
K	13327	0:08:53:08
K	13339	0:08:53:56
K	13398	0:08:55:92
K	13521	0:09:00:84
K	13542	0:09:01:68
K	13574	0:09:02:96
K	13583	0:09:03:32
K	13591	0:09:03:64
K	13597	0:09:03:88
K	13607	0:09:04:28
K	13674	0:09:06:96
K	13693	0:09:07:72
K	13720	0:09:08:80
K	13736	0:09:09:44
K	13760	0:09:10:40
K	13789	0:09:11:56
K	13861	0:09:14:44
K	13883	0:09:15:32
K	13893	0:09:15:72
K	13912	0:09:16:48
K	13919	0:09:16:76
K	13972	0:09:18:88
K	14004	0:09:20:16
K	14017	0:09:20:68
K	14042	0:09:21:68
K	14056	0:09:22:24
K	14060	0:09:22:40
P	14137	0:09:25:48
K	14185	0:09:27:40
K	14196	0:09:27:84
K	14200	0:09:28:00
K	14215	0:09:28:60
K	14219	0:09:28:76
K	14291	0:09:31:64
K	14295	0:09:31:80
K	14330	0:09:33:20
P	14349	0:09:33:96
K	14366	0:09:34:64
K	14407	0:09:36:28
K	14420	0:09:36:80
K	14424	0:09:36:96
K	14434	0:09:37:36
K	14465	0:09:38:60
K	14478	0:09:39:12
K	14533	0:09:41:32
K	14557	0:09:42:28
P	14571	0:09:42:84
K	14584	0:09:43:36
K	14657	0:09:46:28
K	14660	0:09:46:40
K	14697	0:09:47:88
K	14713	0:09:48:52
K	14732	0:09:49:28
K	14746	0:09:49:84
K	14765	0:09:50:60
K	14793	0:09:51:72
K	14804	0:09:52:16
K	14834	0:09:53:36
K	14855	0:09:54:20
K	14886	0:09:55:44
K	14896	0:09:55:84
K	14934	0:09:57:36
K	14983	0:09:59:32
K	15010	0:10:00:40
K	15049	0:10:01:96
K	15099	0:10:03:96
K	15156	0:10:06:24
K	15189	0:10:07:56
K	15211	0:10:08:44
K	15218	0:10:08:72
K	15226	0:10:09:04
K	15237	0:10:09:48
K	15264	0:10:10:56
K	15270	0:10:10:80
K	15279	0:10:11:16
K	15287	0:10:11:48
K	15376	0:10:15:04
K	15379	0:10:15:16
K	15382	0:10:15:28
K	15388	0:10:15:52
K	15396	0:10:15:84

K	15435	0:10:17:40
K	15463	0:10:18:52
K	15470	0:10:18:80
K	15475	0:10:19:00
K	15505	0:10:20:20
K	15530	0:10:21:20
K	15539	0:10:21:56
K	15563	0:10:22:52
K	15610	0:10:24:40
K	15663	0:10:26:52
K	15692	0:10:27:68
K	15704	0:10:28:16
K	15762	0:10:30:48
K	15777	0:10:31:08
K	15848	0:10:33:92
K	15870	0:10:34:80
K	15881	0:10:35:24
K	15893	0:10:35:72
K	15898	0:10:35:92
K	15906	0:10:36:24
K	15911	0:10:36:44
K	15929	0:10:37:16
K	15991	0:10:39:64
K	16001	0:10:40:04
K	16006	0:10:40:24
K	16035	0:10:41:40
K	16048	0:10:41:92
K	16057	0:10:42:28
K	16097	0:10:43:88
K	16164	0:10:46:56
K	16187	0:10:47:48
K	16227	0:10:49:08
K	16288	0:10:51:52
K	16336	0:10:53:44
K	16351	0:10:54:04
K	16375	0:10:55:00
K	16379	0:10:55:16
K	16419	0:10:56:76
K	16425	0:10:57:00
K	16438	0:10:57:52
K	16456	0:10:58:24
K	16487	0:10:59:48

Bibliography

[1] L-P. Bala, K. Talmi and J. Liu.
"Automatic Detection and Tracking of Faces and Facial Features in Video Sequences."
Picture Coding Symposium, 10-12 Sept. 1997, Berlin; Germany.

[2] C. Charayaphan and A. Marble.
"Image Processing system for interpreting motion in American Sign Language."
Journal of Biomedical Engineering, 14:419–425, September 1992.

[3] J. Davis and M. Shah,
"Visual gesture recognition",
IEE Proceedings - Vision, Image, and Signal Processing, Stockholm, May, pp. 101-106, 1994.

[4] T. Starner and A. Pentland,
"Visual recognition of american sign language using hidden markov models,"
In International Workshop on Automatic Face and Gesture Recognition, (Zurich, Switzerland), 1995.

[5] S. Tamura and S. Kawasaki,
"Recognition of sign language motion images",
Pattern recognition, Vol. 21(4), pp. 343–353, 1988.

[6] J.Yang and A. Waibel.
"A Real-Time Face Tracker."
Proceedings of the Third IEEE Workshop on Applications of Computer Vision (Sarasota, Florida, 1996), pp. 142-147.