Efficient Compression for Contour Images

Roumen Kountchev[1], Vladimir Todorov[2], Roumiana Kountcheva[3]

Abstract - **In the paper is presented new method for efficient compression of contour images extracted from video files. The work is aimed at processing of video presentations of Sign Language Interpreter, explaining the training lectures for deaf-impaired students. The presented method offers to use the contour image of the interpreter instead of traditional video as an additional tool included in the distance learning procedure. For this the video is processed frame by frame, the contour image of the interpreter is extracted and the obtained contour images from the consecutive TV frames are compressed and prepared for transfer via Internet. The method for lossless compression of the obtained images is based on image histogram analysis and special Run-Length Coding. In result is obtained new video presentation, which retains the understandability of the training information, but is easily accessed via modem. The efficiency of the new method is higher than that of JPEG2000 (lossless version).**

Keywords **– Lossless image compression, Distance learning, Sign language interpretation.**

I INTRODUCTION

The training of hearing-impaired people based on up-to-date means for video communications is a problem, on which significant experience has been accumulated and many investigations have been carried out [1, 2, 3, 4]. One of the most appropriate means is the Internet, which permits distance learning to be done without live sign language interpretation. The most popular systems used to transfer visual information for distance training for hearing-impaired people via Internet, are based on the video compression standards H.261 [5] and MPEG-4 [6]. In these cases, significant part of the transmitted information comprises the training course, together with video sequences, presenting the image of a Sign Language Interpreter, SLI (Fig.1.). The systems used for the transmission of moving images of this kind [7, 8, 9, 10, 11] in accordance with the already mentioned standards as a rule are based on the principle for inter-frame prediction with movement compensation for every block of the predicted frames of type P. The general disadvantage of these systems is that as a result of the standard compression used to store the presented lesson, the quality of the fast moving objects (in this case, the hands of the sign language interpreter) is significantly deteriorated and the understandability of the presented signs is decreased.

[1] Roumen Kountchev is with the FCTT, Radiocomunications Dept., Technical University of Sofia, Bul. Kl. Ohridsky 8, Sofia, Bulgaria, E-mail: rkountch@tu-sofia.bg

[2,3] Vladimir Todorov and Roumiana Kountcheva are with T&K Engineering, Mladost 3, P.O.Box 12, Sofia 1712, Bulgaria
E-mails: todorov_vl@yahoo.com. kountcheva_r@yahoo.com

Another important disadvantage is that the lessons are presented as video files, which are large and when transferred via modem, the downloading process is relatively slow. In order to avoid the mentioned dynamic distortions, it was found that it is better to compress the video sequences of moving images with intra-frame compression in correspondence with standards M-JPEG [6] or M-JPEG 2000 [12]. In these cases, the image quality is improved, but the compression ratio is low, which results in even larger files being transmitted and received, making the on-line communications more difficult. Similar image deteriorations are obtained in all cases, when the video processing is performed with MPEG-4 [12].

In order to solve the problems mentioned above, in the paper is presented new approach that aims to promote the accessibility to different scientific courses to students with disabilities. Instead of using the full-color video clips, here is proposed to use the contour image of the SL interpreter, which can be compressed and stored with lossless intra-frame compression for future use (Fig. 2).

The Work Hypothesis is based on the investigations presented in works [3,4,7], in accordance with which, for the successful understanding of the presented information is enough to use the contour images of the sign language interpreter only. The basic requirement in this case is the contours to be extracted from the grayscale images (the consecutive TV frames), compressed and after that - transferred with retained quality to the receiving side. The image quality is extremely important for the contours of the face and hands of the sign language interpreter. For this, the intra-frame compression is better than the inter-frame one. The compression of the contour images performed with MJPEG is with low efficiency and the MPEG compression decreases significantly the image quality. The new method solves these problems to a high degree.

Figure 1. Single TV frame with sign language interpreter

Figure 2. Contour image (single frame), 320 x 240 pixels, 8bpp.

The paper is arranged as follows: in Section II is presented the method for lossless compression of still images, in Section III are given the experimental results and Section IV is the Conclusion.

II ALGORITHM FOR LOSSLESS COMPRESSION

The method for lossless compression of still images containing mainly some kind of graphic information, comprises the basic steps, described below:

Step 1. ***Calculation of the image histogram***

The image histogram is calculated in accordance with the traditional approach.

Step 2. ***Histogram analysis***

In result of the histogram analysis is obtained information about:

- The most frequently met values in the processed image (usually these values are zeros).
- The sequences of not used ("free") brightness values in the processed image. The position (the start value) and the length of the longest sequence of not-used values are stored.
- The single not used ("free") brightness values – they are used for the coding only in case that there were not found sequences of not used brightness values in the processed data.
- Sequences of same values in the processed image data and their lengths.

Step 3. ***Coding***

The coding is performed in accordance with several basic rules:

- Every sequence of same values (zeros), whose length is smaller than that of the longest sequence of "free" values is substituted with one of these values, for example: length of two equal values is substituted by the first "free" value from the sequence; length of three equal values – by the second "free" value from the sequence, etc. In result, every such sequence is coded with one byte only;
- Every sequence of same values (zeros), whose length is longer than that of the longest sequence of "free" values, is coded as "value" and "length". These sequences are coded with 2 Bytes code combinations;
- Every sequence of same values (non-zero) is coded as "value" and "length", i.e. - with 2 Bytes code combination.
- Very long sequences of same values (zero or non-zero) are coded as "value" and "length", but using several bytes for the coding.

Step 4. ***Formatting***

The coded images are stored in new special format (tk). For this in the process of coding is created the format header, containing the information about the length of the sequence of "free" values and their start position in the brightness range, together with the most frequent value in the processed image. In the coded image header are used several flags, indicating the way of coding.

The **Block Diagram** of the method is presented in Fig.3.

III EXPERIMENTAL RESULTS

The algorithm, presented above, is very efficient when contour images or graphics are compressed. In order to obtain high efficiency when the video presentations of Sign Language interpreter have to be processed, the video files should be pre-processed properly. The pre-processing comprises the following two steps:

- The original video sequence (which had not been compressed with H.261 or MPEG-4) is stored in digital form in order to make every TV frame available for processing;
- Each TV frame is processed individually. The processing comprises:
  - Converting to grayscale;
  - Extraction of the main contours.

In result is obtained the contour image of the Sign Language Interpreter. After these procedures the image is ready for the compression.

Some of the obtained experimental results are presented below. In Fig. 4 are shown four contour images, obtained from corresponding consecutive TV frames.

Fig. 4. Test contour images, obtained from consecutive TV frames: 01c, 02c, 03c, 04c.

Image	01c	02c	03c	04c	05c	06c	07c	08c
TKView	38.9	38.9	39.02	38.5	39.02	39.28	39.16	38.7
JPEG	3.34	3.35	3.36	3.36	3.30	3.35	3.37	3.37
JPEG2000	5,82	5,81	5,82	5.80	5.81	5.83	5.77	5.77

Table 1. Compression ratios, obtained with the new method for intra-frame compression and JPEG standards.

The presented results were obtained with the program TKView, developed on the basis of the described algorithm; the results for the JPEG standard were obtained with Microsoft Photo Editor (highest quality, lossy compression), and for JPEG2000 – with Algovision LuraTech (lossless compression).
The mean compression ratio obtained for 90 test images with TKView is 39,08; with JPEG (best quality) it is 3,36 and for JPEG2000 lossless – 5,80. For the tested kind of images the results obtained for the three methods are very consistent.
The comparison results show that the new method is 11,63 times more efficient than JPEG (MS Photo Editor, best quality) and 6,74 times more efficient than JPEG2000 (lossless) for the investigated image class.

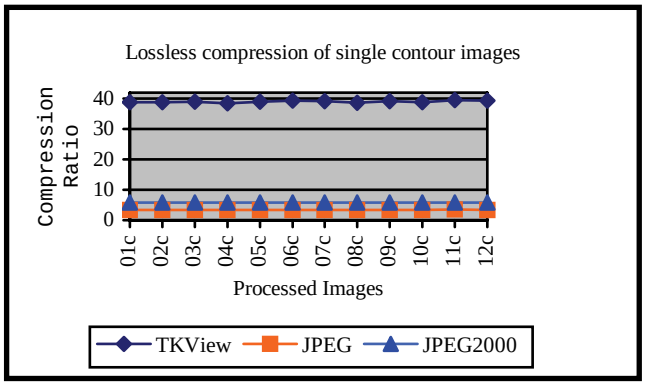

Fig. 5. Graphic presentation of the results obtained with inter-frame compression of contour images

In order to evaluate the method efficiency, the following is assumed: Each TV frame from the original video lesson, representing the greyscale contour image of the SL interpreter is treated as a single still image (320 x 240 pixels) with size 76800 bytes. After compression accordingly the new method (about 40 times) with the special software for lossless compression, the size of the file, obtained in result, is less than 2000 Bytes. The single images obtained from the corresponding consecutive TV frames after that are arranged as video clips and transferred to the receiving side, where they are restored. The high compression ratio permits relatively easy to add such information to distance-learning lessons, accessed via Internet.

IV Conclusion

The presented results show consistent high compression ratios when contour images are processed. These compressions are significantly higher than those obtained with software, based on the JPEG2000 standard (lossless).

In accordance with the presented results, the contour image of the sign language interpreter obtained from a single TV frame with size 320 x 240 pixels after lossless compression is less than 2 KB. This means that the compressed video data could be transferred as additional information, accomplishing the distance-learning lectures for hearing-impaired students. For presentations of 25 TV frames per second the necessary additional information is about 49 KBps. As it is known, the standard MPEG-2 bit-rate is in the range 6-12Mbps. The additional information required for the presentation of the sign language interpretation will require the insertion of these additional 49 KBps.
In case that the size of the used TV frame is smaller (for example, 256x192 pixels) the mean compression ratio is 38 and the additional bit rate is about 30 KBps.
The presented approach is suitable for applications concerning distance learning for students who are deaf or hearing impaired. It could be used for the creation of special handbooks and vocabularies in all cases when some kind of additional information is required and the sign language interpretation will make the explanation easier.

Another possible application is the compressed information to be inserted in some multimedia products like DVDs in order to provide the required information to the hearing impaired people in their most natural way. Same approach could be used in the TV broadcastings instead of the image of the live sign language interpreter. For these applications the contour image could be presented in a small window placed in the lower part of the TV screen.
The comparison results with the JPEG2000 standard for lossless compression of contour images show the advantages of the presented method.
The method is suitable for efficient compression of text and other graphic images, which additionally supports all kinds of distance learning applications [13].

The **further development** of the method is aimed at the following directions:

- Development of special algorithms for efficient extraction of the contours of the interpreter's hands and face in order to optimize the presented information, retaining its understandability.
- Investigation on the influence of the participating brightness levels in the extracted contours on the compression ratio.
- Investigation on the influence of the number of consecutive TV frames, compressed together, on the method efficiency.
- Adaptation of the compressed data header with the requirements of broadcasting standards and DVD formats.
- Development of a special version of the method, aimed at services in mobile communications.

V Acknowledgement

This paper was supported by NSF of Bulgarian Ministry of Education and Science (Contract VU-MI-104).

REFERENCES

[1] Subject-Based Deaf and Hard of Hearing Internet Resources. www.Internet Resouces - Deafness.htm.
[2] Teaching the Learning Disabled: Study Skills and Learning Strategies. www.Teaching the learning disabled.htm.
[3] D. Pearson, J. Robinson. Visual Communications at Very Low Data Rates. IEEE Proc., Vol. 73, 1985, pp. 222-241.
[4] P. Letellier, M. Nadler, J. F. Abramatic. The Telesign Project. Proc. of the IEEE, Vol. 73, 1985, pp. 242-258.
[5] CCITT Recommendation H.261: Video Codec for Audio- visual Services at px64 kbit/s.
[6] M. Ghanbari. Standard Codecs: Image compression to Advanced Video Coding. Institution of Electrical Engineers, 2003.
[7] M. Whybray, E. Hanna. A DSP based videophone for the hearing impaired using valledge processed pictures. Proc ICASSP'89, Vol 3, 1989, pp.1866-1869.
[8] M. Whybray. Moving picture transmission at low bitrates for sign language communication. http://www.zak.co.il/deaf-info/old/videophones.html.
[9] G. Sperling, M. Landy, Y. Cohen, M. Pavel. Intelligible encoding of ASL image sequences at extremely low information rates. Computer Vision, Graphics, and Image Processing 31, 1985, pp. 335-391.
[10] S. Tetzchner (Ed.). COST 219 book: Issues in Telecommunications and Disability. Commission of the European Communities, Catalogue number CD-NA-13845-EB-C, 1991.
[11] A. Heron, N. Mc Donald. Video transmission over a radio link using H.261 and DECT. 4-th International Conference on Image Processing and Applications, IEE, 7-9 April 1992.
[12] D. Salomon. Data compression. Springer-Verlag, New York, Inc. 2004.
[13] Kountchev R., Todorov Vl., Milanova M., Kountcheva R. *Documents Image Compression with IDP and Adaptive RLE*. The 32nd Annual Conference of the IEEE Industrial Electronics Society Conservatoire National des Arts & Métiers – Paris - FRANCE, IECON'06, November 7 – 10, 2006, pp.3261-3266.

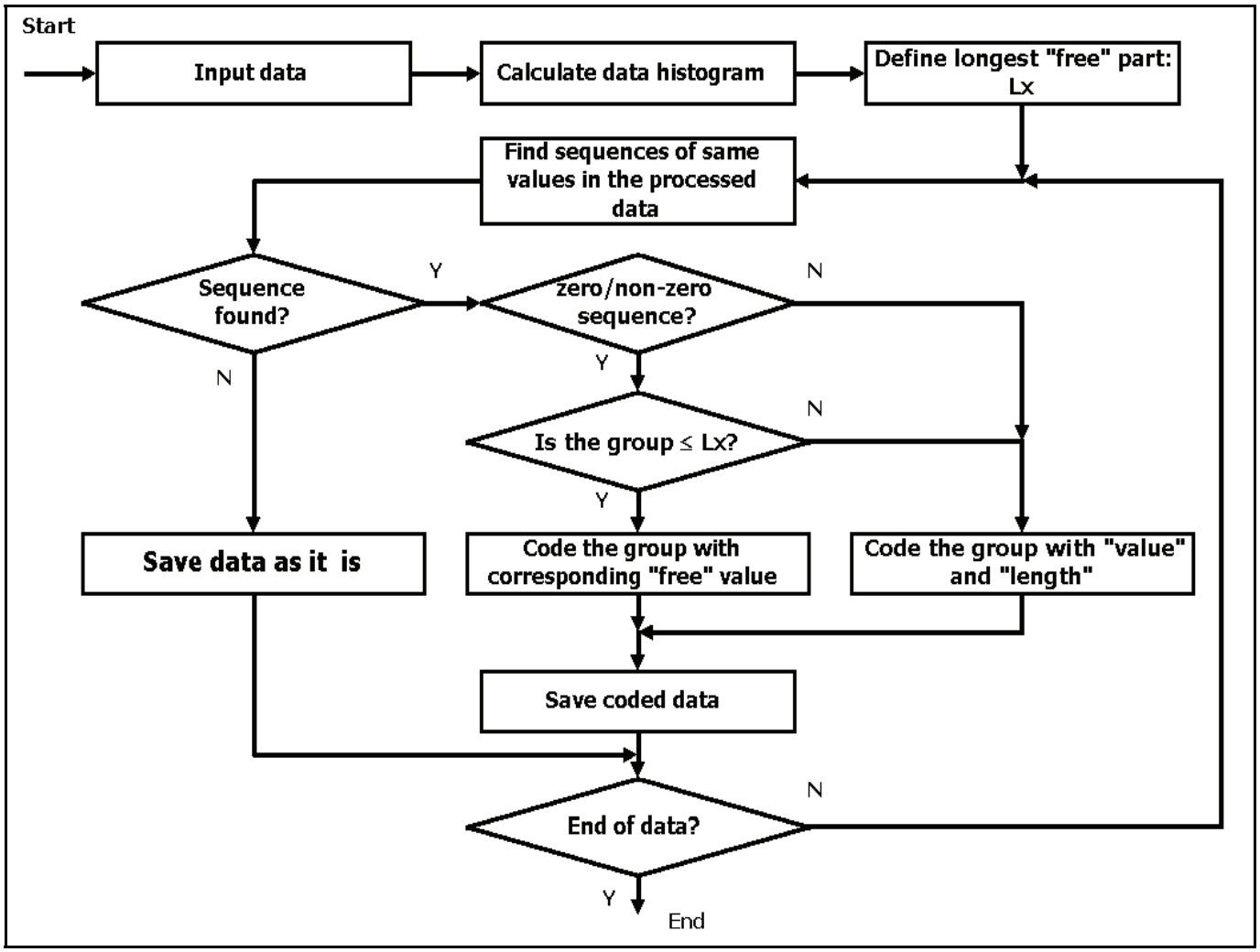

Fig. 3. Block diagram of the method for lossless coding of still images