Processing of Sequences of Contour Images of Sign-Language Interpreter

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

Abstract - **In the paper is offered a new approach used for the processing of sequences of TV frames, representing the performance of a sign language interpreter. The processing comprises two steps: the extraction of the image contours and their processing with intra-frame lossless compression aimed at retaining the understandability of the presented signs. The efficiency of the lossless compression is increased when the extracted contour images are processed in groups, comprising several consecutive TV frames. The obtained high compression ratio permits the application of the method for distance learning applications in training courses for impaired students and in some real-time applications.**

Keywords – **Lossless image compression, Distance learning tools**

I. Introduction

The distance learning for hearing impaired people is a problem of high significance. The usual approach is to provide them the training courses as pdf or doc files via Internet, or to use video presentations, containing explanations of sign language interpreter (SLI). The approach based on the participation of SLI is of great importance because the trainees assess the information using their natural language. Insufficiency of the method is that such video presentations offer very bad quality of the fast moving parts such as the interpreter's hands and face, which are of great importance for the signs understandability and the transferred volumes of data are very big. In some cases instead of the color video presentations are used the contours of the SLI image [1,2,6]. For this, the contours are extracted from the video frames and after that - compressed. As it is presented in the special literature [3,4,5], the contour images are understandable and present the information in a quick and acceptable way. The main problem remains the large amount of data, which should be transferred and received [11,12]. The efficient coding of this information will permit its transfer via slow-rate channels [7,8,9,10]. The investigation, presented in this paper is aimed at the efficient compression of the contour images. For this, the consecutive TV frames are processed one by one in order to extract the contours of the SLI image and to prepare them for the compression.

[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

Usually there are too many extracted contours, which is an obstacle for the efficient compression. In order to preserve the main contours only, retaining the signs understandability, in this work is used presented special pre-processing, based on the image histogram modification.

The paper is arranged as follows: in Section II is presented the algorithm for image pre-processing, in Section III are given the experimental results, obtained for single and multiple frame compression and is investigated the influence of the number of consecutive TV frames processed together; Section IV is the Conclusion.

II. Image pre-processing

The method for image pre-processing performs adaptive image contrast enhancement, comprising two consecutive stages: brightness segmentation based on the image histogram analysis, and transformation of the pixels' brightness in accordance with tables, defined by the segments treatment.

In the first stage is performed the image segmentation. For this are used the thresholds k_1 and k_2, which divide the histogram in three segments (A, B, C). The thresholds are set so that to define the second segment (B), which contains the main part of the image objects. In order to make the number of participating brightness levels in the segment B smaller (and correspondingly - to obtain higher compression), this part should be skewed to 1-2 % (instead of 12). The histogram of the SLI image (Fig.1.) is shown in Fig.2 together with the histogram segmentation.

Fig.1. TV frame with SLI

The limits of the central segment (B) are defined performing the following operations:

- The image histogram h(k) is calculated and is defined its maximum:

$$h_{max} = \max\{h(k)\} \text{ for } k = 0, 1, 2, \ldots, k_{max},$$

- The value $t = \alpha\, h_{max}$ is defined, for $\alpha < 1$ (for example $\alpha = 0.8$): this value defines the magnitude of the middle segment of the histogram (B).

- The values of the start and end points of the segment B, defined as k_1 and k_2, are calculated in accordance with the relations:

$h(k) \leq t$ for $k = 0, 1, 2, \ldots, k_1\text{-}1$,

$h(k) \geq t$ for $k = k_2+1, k_2+2,.., k_{max}$, for $|k_2 - k_1| \geq \Delta$.

(Δ – a value, set in advance).

In case, that the last condition is not satisfied, the value of the parameter α is decreased, for example to α=0.75, and using the new threshold value $t=\alpha.h_{max}$ are defined the corresponding values for k_1 and k_2. When the requirement $|k_2-k_1| \geq \Delta$ is satisfied, the calculation cycle ends; if not – the process continues. An example is presented in Fig.2.

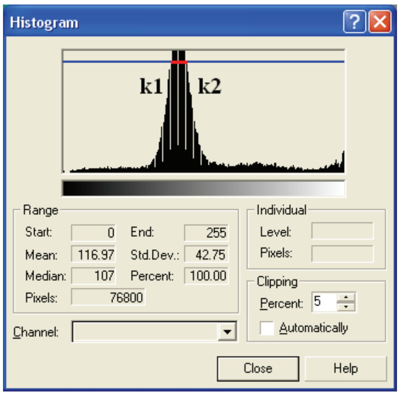

Fig.2.The histogram of the SLI image: the three segments of the image histogram, defined by the points k_1 and k_2

In the second stage of the processing, the brightness level k of the pixels in the three segments of the histogram is transformed in accordance with the relations:

$$g(k) = \begin{cases} g_A(k) & \text{if} \quad 0 \leq k < k_1; \\ g_B(k) & \text{if} \quad k_1 \leq k \leq k_2; \\ g_C(k) & \text{if} \quad k_2 < k \leq k_{max}. \end{cases}$$

The relations $g_A(k)$, $g_B(k)$ and $g_C(k)$ represent the brightness transformation for the pixels in the segments A, B and C correspondingly. In order to perform the required contrast modification, the boundaries k_1, k_2 of the segment (B) are skewed to $(k_1+\delta_1)$ and $(k_2-\delta_2)$, and correspondingly are moved the upper limit value of the segment A and the lower limit value of the segment C. In the presented example δ_1 and δ_2 are parameters, which define the contrast modification of the objects in the segment B and as a consequence – of the segments A and C as well. The brightness transformation tables are defined in accordance with the requirement for histogram equalization of the corresponding segment (A, B or C) with changed (stretched or skewed) brightness range:

$$g_A(k) = (k_1 + \delta_1)\sum_{l=0}^{k} h_A(l),$$

$$g_B(k) = (k_2 - k_1 - \delta_1 - \delta_2)\sum_{l=k_1+\delta_1}^{k} h_B(l) + (k_1+\delta_1),$$

$$g_C(k) = (k_{max} - k_2 + \delta_2)\sum_{l=k_2-\delta_2}^{k} h_C(l) + (k_2 - \delta_2).$$

In particular, for images in which the histogram of the corresponding segment is uniform, i.e. for:

$$h_A(k) = \frac{1}{k_1} \quad \text{for} \quad k = 0,1,..,k_1-1;$$

$$h_B(k) = \frac{1}{k_2 - k_1} \quad \text{for} \quad k = k_1, k_1+1,.., k_2;$$

$$h_C(k) = \frac{1}{k_{max} - k_2} \quad \text{for } k = k_2+1, k_2+2,.., k_{max},$$

the relations for the brightness transformation are linear and are defined as follows:

$$g_A(k) = \left(\frac{k_1 + \delta_1}{k_1}\right) k;$$

$$g_B(k) = \left(\frac{k_2 - k_1 - \delta_1 - \delta_2}{k_2 - k_1}\right)(k - k_1) + (k_1 + \delta_1);$$

$$g_C(k) = \left(\frac{k_{max} - k_2 + \delta_2}{k_{max} - k_2}\right)(k - k_2 - 1) + (k_2 - \delta_2 + 1).$$

In this case the brightness levels in the range (k_1, k_2) are skewed and correspondingly - the brightness levels in (0, $k_1 + \delta_1$) and ($k_2 - \delta_2$, k_{max}) are stretched.

III. EXPERIMENTAL RESULTS

A. Single frame compression

In result of the performed histogram modification the number of the participating brightness levels is decreased, which results in higher compression ratio. The histogram modification limits the number of the retained brightness levels to less than 5 in the meaning part of the image. In the remaining part (the image background) the number of the brightness levels is very low too. In result of the performed histogram modification the compression ratio is significantly increased. For the investigation was used more than 120 test images extracted from different video sequences. The results obtained for contour images extracted from different sequences are very close and consistent. Some of the results are presented in Table 1 below. The mean value for the compression ratio after the processing is more than 15 times higher than that, obtained for the original image.

Image	C1	C2	C3	C4
Original	2,54	2,58	2,29	2,49
Treated	36,56	36,78	36,79	37,01

Table 1. Compression ratio obtained with the new method for image processing and lossless compression.

B. Multi-frame compression

In order to obtain higher compression ratio is used multi-frame lossless compression of the contour SLI images extracted from consecutive TV frames. This approach aims to use the high correlation between the consecutive TV frames. The investigation was performed, arranging the processed contour images in one larger image with different size: 2, 3, … up to 12 consecutive frames, arranged horizontally, vertically or as a rectangle. The obtained results are presented below.

Arrangement in horizontal direction

The multi-frame lossless compression is performed arranging the images together and processing the thus obtained compound image as a single one. Here L1h (with size 320 x 240 pixels) is the contour image obtained after contour extraction and histogram modification of the first TV frame in the processed sequence; L2h comprises the contour images obtained from two consecutive TV frames and is with size 640 x 240 pixels; L3h comprises the contour images obtained from three consecutive TV frames and is with size 960 x 240 pixels, etc. The last test image, L9h comprises the contour images obtained from nine consecutive TV frames. Example test images are shown in Fig. 3.

Fig.3. a. Image L1h.

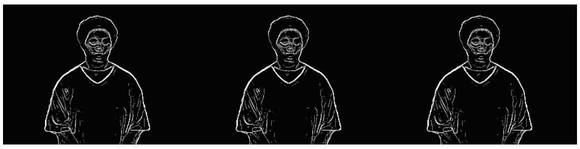

Fig.3. b. Image L3h.

Image	L1h	L2h	L3h	L5h	L6h	L8h	L9h
TKView	40,02	40,39	40,53	40,54	40,62	40,65	40,70
JPEGmax	3,34	3,35	3,36	3,38	3,39	3,39	3,39
JPEG2000	5,82	5,96	5,94	5,97	5,98	5,98	5,99

TABLE 2. Compression ratios for inter-frame compression with horizontal arrangement

The obtained results are presented in the Table 2 and in Fig.4. The last image, L9h comprises the contour images extracted from 9 consecutive TV frames. The comparison was performed with software products, based on the standards JPEG and JPEG2000, and the special lossless compression was performed with TKView. The results, obtained with TKView are much better than the other two. The results for JPEG are not for lossless compression, but for lossy compression with highest possible quality (Microsift Photo Editor, quality factor 100).

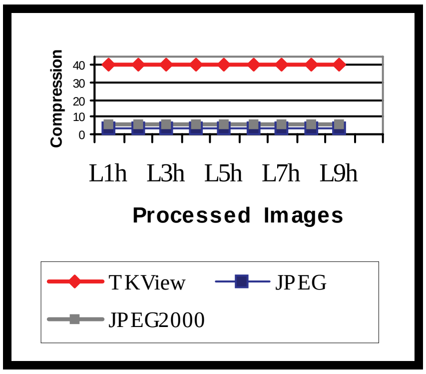

Fig.4. Compression ratio, obtained with horizontal arrangement of the tested images.

Arrangement in vertical direction

Similar investigation was performed with vertical arrangement of the tested images and processing them together. The image L1v is the contour image obtained from the first TV frame in the sequence, with size 320x240 pixels; L2v comprises the contour images obtained from two consecutive frames and is with size 320 x 480 pixels; L3v comprises the contour images obtained from three consecutive TV frames and its size is correspondingly 320 x 720 pixels, etc. The last image, L9v comprises 9 contour images obtained from consecutive TV frames. Some of the obtained results are given in the Table 3.

Image	L1v	L2v	L3v	L5v	L6v	L8v	L9v
TKView	40,02	40,16	40,27	40,24	40,33	40,35	40,39
JPEGmax	3,34	3,37	3,38	3,38	3,39	3,39	3,40
JPEG2000	5,82	5,90	5,91	5,92	5,93	5,94	5,94

TABLE 3. Compression ratios for inter-frame compression with vertical arrangement

The mean compression ratio for TKView is 40,25; for JPEG (lossy compression with highest quality) it is 3,38 and for JPEG2000 – 5,91. The results, obtained for a big number of test images are very consistent. The results show that the compression ratio is increased together with the number of comprising images from 1 to 9; after that the change is negligible. In Fig. 5 the obtained results are presented graphically.

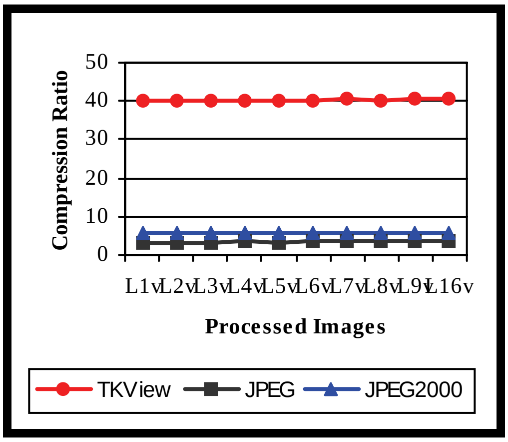

Fig. 5. Graphic presentation of the compression ratio obtained with vertical arrangement

2D Arrangement

The 2D arrangement of the tested images was performed to investigate the influence of the correlation in both directions (vertical and horizontal) on the compression ratio. The processed images were arranged as follows: the image L2x3 comprises 6 consecutive contour images arranged in two rows of 3 images each; the image L2x4 comprises 8 consecutive contour images arranged in two rows of 4 images each, etc. The test image L4x4 is with size 1280 x 960 pixels. The results obtained for the compression ratio are given in Table 4.

Image	L2x3	L2x4	L2x5	L3x4	L4x3	L3x3	L4x4
TKView	40,90	40,98	40,99	41,02	41,06	41,12	41,14
JPEGmax	3,39	3,36	3,30	3,32	3,31	3,32	3,33
JPEG2000	5,92	5,92	5,93	5,93	5,94	5,94	5,94

TABLE 4. Compression ratios obtained with 2D arrangement

The mean value for the compression ratio obtained with 2D arrangement for many test images is as follows: for TKView it is 41.01; for JPEG (lossy compression with highest quality) it is 3.33 and for JPEG2000 (lossless) – 5.93.

The results, obtained with the presented three kinds of arrangement show that highest compression is obtained for 2D arrangement of the tested images. The optimal number of the images processed together is 12. The comparison results with JPEG and JPEG2000 are shown in Fig. 6.

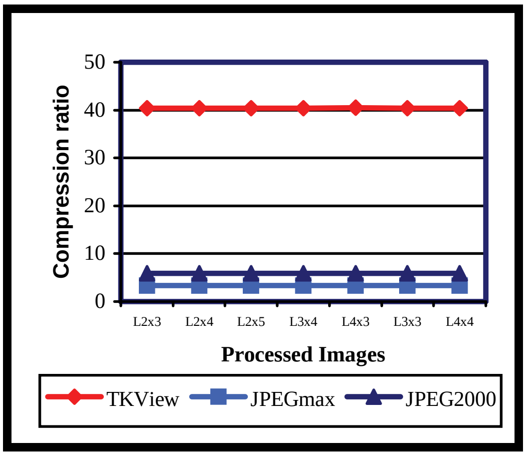

Fig. 6. Graphic presentation of the compression ratio obtained with 2D arrangement

Arrangement	Mean compression Ratio
Horizontal	40.65
Vertical	40.85
2D	41.01

TABLE 5. Mean values of the obtained compression ratio

In Table 5 are presented the mean values for the compression ratio obtained with different arrangement for same images.

IV. CONCLUSION

The presented method for image histogram modification influences the compression ratio of contour images to a high degree. Very important quality of the method is that the understandability of the presented signs is retained because the TV frames are processed individually and after that - compressed. The method permits the compressed data to be transferred as additional information via low-rate channels. The compression results are much better than those, obtained with traditional methods based on the JPEG and JPEG2000 standards. The specific features of the method are:

- The efficiency of the method depends on the retained number of brightness levels in the extracted contour images;
- The consistency of the obtained results permits the parameters, used for the histogram modification to be selected in advance and to perform the pre-processing automatically, which is very important for real-time applications.
- The consecutive TV frames could be arranged after decompression and used as usual video clip.
- In result of the high compression ratio one TV frame, containing contour SLI image is compressed to less than 1900 KB, i.e. the additional information, needed for the video SLI presentation is less than 380 Kbps. Compared with the regular data flow in MPEG-2, which is in the range 6 - 12 Mbps, this amount of data is negligible.
- Additional possibility is to use TV frame with smaller size, for example 256 x 192 pixels. In this case the mean compression ratio is 38 and the additional data flow is about 260 Kbps.
- The method for lossless compression gives very good results when applied for text images and this is an additional tool for distance learning applications [13].

The presented approach is suitable for various applications aimed at training of hearing impaired students.

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] 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. IEE, 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, 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, April 1992.

[12] D. Salomon. Data compression. Springer-Verlag, New York, Inc. 2004.

[13] R. Kountchev, Vl. Todorov, M. Milanova, R. Kountcheva. *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. 3261-326.