

Synthesizing Intermediate Images Using Stereoscopic Images

Jihyun Kwak*, V. S. V. Komar, and Kyung-Tae Kim

*Dept. of Information & Communication Engineering, Hannam University,
Daejeon 306-791, KOREA*

(Received August 8, 2002)

In this paper, we present an algorithm for synthesizing intermediate views from a stereoscopic pair of images. Syntheses of intermediate images allows one to realize a more comfortable the 3D display system. The proposed method is based on block matching, which is not ordinarily used. The contour information is used for a block decision. In order to find an equivalent (or corresponding) block, there are two steps: "matching of contour-to-original image" and "matching of contour-to-contour image" methods. "Matching of contour-to-contour image" uses both left and right contour images. This block matching method allows us to find the corresponding block in spite of different block sizes. Experimental results illustrate the performance of the proposed technique and we obtained a high quality image of more than 31 dB PSNR.

OCIS code : 100.6890.

I. INTRODUCTION

Since the middle of the 19th century 3D display systems have caught human beings' interests, and they are on the way to being fully developed. 3D display systems can be used in communication systems, education, medical science, and commercial advertisement [1]. To be able to develop good 3D display systems, two conditions must be fulfilled: first, systems need to be realized by the current hardware capacity, and second, the systems need to be made familiar to the human beings who are not familiar with seeing 3D images. However, to be able to represent 3D images, two images (left and right images) at least are required, which brings the systems to have wider bandwidths. Consequently, it requires larger and faster memory, network, and computer systems. Moreover, 3D display systems must not make people tired, and they need to have wide viewing zones to avoid discomfort [2].

To solve most of these problems, many researchers including ourselves have worked on synthesizing intermediate images [3-7]. But the quality of synthesized images produced by most of researchers is not so high [9,10]. Synthesizing intermediate images can be defined as follows: by using two stereo images, an intermediate image can be synthesized at an arbitrary point between two stereo images. For synthesized in-

termediate images to be realized, the systems should be simplified and less data should be sent and saved.

In this paper, original images are changed to contour images by the threshold method, blocks between two adjacent contours are divided by the segmentation method and matched points are found by comparing left and right images to synthesize an intermediate image. Also, this paper describes how shadow blocks can be processed. The proposed method uses two kinds of matching algorithms: first, a block matching method which is ordinarily used between block and original images is used; and second, a block matching method between two blocks of left and right images is used. This method is based on the fact that corresponding block sizes of left and right images are different when the objects incline to the depth direction. Namely, this proposed method can find different sizes of equivalent blocks of both images between left and right images.

This paper is formatted as follows: the second section describes briefly the whole algorithm to synthesize intermediate images and choose the blocks by contour images; the next step includes the decision regarding equivalent blocks and shadow blocks, and how they can be processed; and the third section describes synthesizing intermediate images using the proposed algorithm, and shows the results by using the proposed method; and finally, conclusions follow.

II. ALGORITHMS

Fig. 1 shows a schematic diagram for synthesizing an intermediate image from stereoscopic images. The intermediate images are synthesized from some regions of real initial images by their preprocessing. The original images of extreme views are required to synthesize the intermediate images. These intermediate images can be introduced as ones which are shot by some virtual cameras at arbitrary views between left and right cameras (extreme views). To synthesize intermediate images, matched corresponding points between left and right original images are needed.

Original images are pre-processed by histogram equalization in each image, with the same average and variance. The next step of processing is to build contour images. Contour images are very important as a preprocess for block decisions. They are used to find equivalent blocks in the next step. In this paper, contour images are built by segmentation of images.

The next step is to find equivalent blocks. Two steps are used for this purpose. The first step is "matching of contour-to-original image" between left and right images, which is a very traditional matching method. It is defined as between a contour image and an initial original image. The second step to find an equivalent block is "matching of contour-to-contour image". It is defined as between as contour image and a contour image. As mentioned above, this matching method is based on the fact that corresponding block sizes of left and right image are different when the objects incline

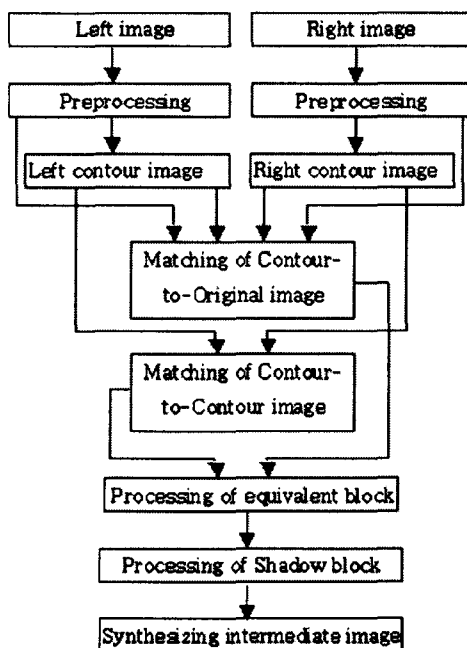


FIG. 1. Schematic diagram for synthesizing the intermediate image.

to the depth direction. Namely, this proposed method can find the different size of equivalent blocks of both images between left and right images.

Finally, all equivalent blocks are found and recorded in one array with its attributes: block type, beginning and end coordinates in left and right selected blocks, and likelihood of equivalence. Arrays are sorted in ascending order according to the likelihood information [3]. Then, equivalent blocks of the intermediate image are made from array information. When all equivalent blocks are decided, empty regions of image are defined as shadow blocks. Shadow blocks are corresponding parts of a scene shot by one camera only. Shadow blocks of intermediate image are filled using the disparity information.

Fig. 2 shows three different regions when an image is taken by two cameras. The first region (1) is called an equivalent block, and the image is seen by both cameras. Matched points can be easily found. The objects (1) incline to the depth direction, therefore it is very natural that the size of image of the objects (1) is different. The second region (2) is called a shadow block, the image is seen by only one camera (either the left camera or right camera, not both). This region needs to be processed with a special method which is explained in section 5. Lastly the third region (3) is the part of the image which is taken by neither the left nor right camera. This region is called a double shadow block [3,4].

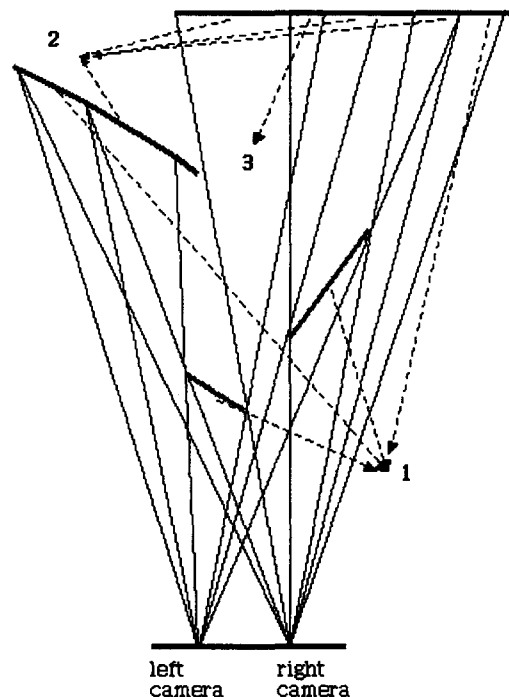


FIG. 2. The three different regions when an image is taken by two cameras.

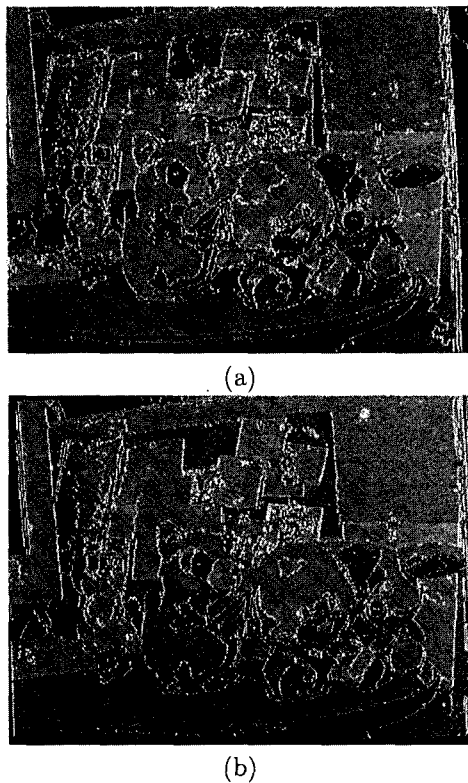


FIG. 3. (a) Left Contour Image. (b) Right Contour Image.

After contour images are built, to define the kinds of contour images is the important preprocess for a block decision. It is used to find an equivalent block which is the next step. In this paper, contour images are made by the image segmentation method. A color difference image is calculated until the differences are less than two thresholds while an arbitrary point moves along the line to be segmented. A color difference no more than threshold 1 for small area and no more than threshold 2 for wide area is controlled. This is one segment. If the greatest point difference in color occurs, then this is an object border (not a flat but object contour). Typical values of thresholds are 15 and 45 units (units stand for 1 of the 255 scale). This step is the end of the first operation (color segmentation). Processing from segmentation to find disparity is repeated several times as defined by the initial condition of the system. From the second time of processing, because we know the disparity information estimated roughly by the previous step, we can use those. Borders of preliminary selected regions are checked by a comparison of the disparity between the inside and outside of the segment (supposed object). Thus selection of objects is made by merging the small objects which have similar disparities with a certain limit. Now borders of segments are borders of objects. Fig. 3 shows the results of segmentation and the original images

together.

1. Finding an equivalent block

Purpose of this step is to decide all blocks which are candidates for a pair of equivalent blocks. It is necessary to save all possible candidates. Equivalent blocks of images are blocks which correspond to some parts of a scene shot by both cameras (region (1) in fig. 2). Two steps are used for this purpose. The first step is a “matching of contour-to-original image” method, which is to find the equivalent block between the left contour image and right original image, then between the right contour image and left original image. The second step of finding an equivalent block is the “matching of contour-to-contour image” method, which is to find the equivalent blocks between the left contour image and right contour image. Equivalent blocks of the left and right images can have different lengths, but the length is not considered in the first step. Therefore, this step takes care of the different lengths of blocks. A longer block is compressed so that its length becomes the length of the smaller block. Then, pixel by pixel comparison can be done. Let us examine each method in depth.

Likelihood is calculated from the elements of difference of length deviation, long block priority, contour contact coefficient, color difference and color ratio [8].

A. Matching of contour-to-original image

Above all, blocks between contour points in every line, are defined as one of the following: equivalent, shadow, or double shadow blocks.

A block which is to be processed at the left contour image is compared with the right image on the same line (see fig. 4). Typical limits of search are ± 50 pixels. Block of the same length move from left to right with a step which can be set as 1 to 5 pixels. The running time of the process depends on this step value significantly. However, if the step value is more

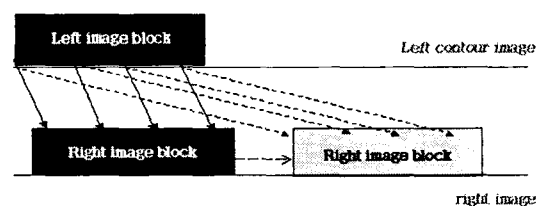


FIG. 4. “Matching of Contour-to-Original image” method.

than 3pixels, the quality of image becomes worse. It has been found experimentally that the optimum step value is 3 pixels.

As a block on the contour image moves along the other original image by the specified step size, likelihood values are calculated. If the calculated likelihood value is less than a permissible value, it resets to zero. Then, a block with the largest likelihood value is selected; and if its likelihood is a non-zero value, the selected block and the block from the contour image are marked as candidates for equivalent blocks. Afterwards the next block from the left contour image is compared with the right image using the same method. When all the first line of the left contour image is compared with the right image, the direction of comparison is changed to a comparison between the first line of the right contour and left contour images. All blocks on the current line are considered. All selected pairs of equivalent blocks are stored. This is the end of the "matching of contour-to-original image" step.

But it is possible that each block of the pair is terminated between contour points. Such blocks can have equal and different lengths. Those blocks are considered next.

B. Matching of contour-to-contour image

At this step all blocks, which are concluded between contour points in the left image, are compared with all blocks between contours in the right image in certain limits (same as during 'matching of contour-to-original image'). Every block in the left image is compared with every block lying on the same line in the right image (see fig. 5). Processing is done in each line from left to right and from top to bottom. The end of each block is the beginning of the next block. The blocks are compared in all pair found. First, blocks are compared by length. The longer block is compressed, so its length becomes the size of the smaller block. Then pixel-by-pixel comparison is done. If the value of equivalence likelihood is not zero, then pairs of the block are stored. A likelihood value that is summed with absolute differences between pixels and normalized with the length of the block is saved. Now

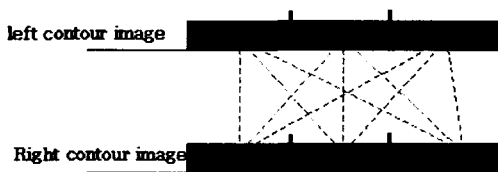


FIG. 5. "Matching of Contour-to-Contour image" method.

there are blocks after the 'matching of contour-to-original image' step and blocks after the 'matching of contour-to-contour image' step as candidates for an equivalent block pair. All these blocks are loaded to one array with the results of the likelihood value of each block.

2. Finding shadow blocks

All information of pairs of equivalent blocks are saved in corresponding arrays. Empty regions excluding the equivalent blocks have to be processed. The borders of equivalent blocks cannot always be defined correctly. The biggest error quantity occurs at the borders of equivalent and shadow blocks. In order to exclude possible errors of definition of borders of a shadow block, the shadow block and adjacent equivalent block are considered as entire blocks; we call this "mixed blocks". Fig. 6 (a) shows the mixed block in the left image, and (b) in the right image.

Either side of a shadow block has a border with an equivalent block which corresponds to a closer object. This border can be a contour point. It is defined exactly in most cases, but other borders are defined in a harder way. There is no contour on the border of

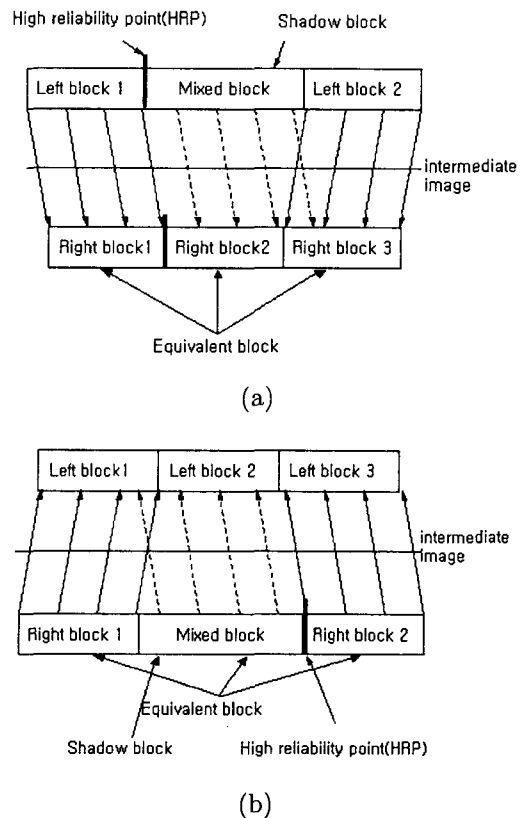


FIG. 6. (a) The mixed block on the left image. (b) The mixed block on the right image.

blocks. Often, this region, which corresponds to a more distant object, has no feature points, color change, etc. By the geometry of viewing, when the left image has shadow blocks, the blocks of intermediate image are filled with the left part of blocks of the left image. When the right image has shadow blocks, the blocks of intermediate image are filled with the right part of blocks of the right image [4]. Therefore, the left border of the left mixed block and the right border of the right mixed block are high reliability points. High reliability points have to be defined correctly with high likelihood. A left mixed block shifts as an equivalent block from the left side. Disparity can be measured for an equivalent block only. As a result, the disparity of a shadow block is forced to have the same disparity of the adjacent equivalent block. This assumption is applied to the right mixed block, too. The right mixed block shifts as an equivalent block from the right side. Only a part of a mixed block will be shifted to an intermediate image after overlaying an equivalent block corresponding to a closer object. This is the part of the shadow block which is hidden in the mixed block. As a result, certain parts of an entire block which correspond to the shadow region and the adjacent equivalent region of an object are synthesized to all required intermediate images. Thus use of mixed block allows us to get a continuous part of an image on the border of an equivalent and shadow block.

After all equivalent blocks in mixed blocks are decided, remaining empty regions of an image are defined as shadow blocks. Empty regions are searched along the processing line from left to right. If the left side of an empty region's disparity is less than the right side of an empty region's disparity, then there is a left shadow block in the left image. It has to be confirmed as a longer empty region in the left image. The preceding equivalent block is merged with a part of the original image which corresponds to an empty region. If the left border of a connected equivalent block is the contour point, then the merged block is a mixed block which is suitable for building intermediate images. If it is not the contour point, the adjacent left equivalent block is connected. In this operation, equivalent blocks of minimum length must be specified. Thus, the left border of a mixed block is a point of high reliability such that there is high likelihood of a correct definition of the left border of a mixed block. Hence the shift of a mixed block in intermediate images will be correct, too. Although the border between an equivalent and shadow block is not defined right, shift is achieved. Moreover, possible distortions on this border are excluded because the mixed block is the entire region of the initial image.

If the left side of an empty region's disparity is greater than the right side of an empty region's dispar-

ity, then there is a shadow block in the right image. It has to be confirmed as a longer empty region in the right image. In this case, the following equivalent block is merged with part of the original image which corresponds to an empty region. If the right border of the connected equivalent block is a contour point, the merged block is a mixed block which is suitable for building intermediate images. If it is not a contour point, the following equivalent block is connected [5-7]. Thus the right border of an equivalent block is a point of high reliability. As in the preceding case of the left shadow block, the correct shift of a mixed block is ensured in intermediate images. Errors on the border of the shadow and the equivalent blocks are excluded also.

III. EXPERIMENTS AND RESULTS

1. Synthesizing intermediate images

All defined blocks are loaded to some arrays. Each block in an array has numbers and a set of attributes. These attributes are: kinds of blocks, beginning point coordinate, end point coordinate, and value of equivalence likelihood.

The program allows us to synthesize an intermediate image for any view with a 1% step (view number 0 - 100). Extreme images which are divided into blocks can be synthesized as any other view image. For this result it is necessary to give view number 0 or view number 100. When there is a request for a certain view image, a block is ordered according to its likelihood values saved in arrays. Then only mixed blocks are selected by a whole array search. Attributes of mixed blocks are: left or right block, coordinates of high reliability point (beginning coordinates in left and right images for left blocks, or end coordinates in the left and the right images for right blocks). The likelihood of mixed blocks is a small formal value, and it is not considered. The most important value is the high reliability point, which defines the shift of the block to any intermediate image. The shift of the whole block is the same as the shift of the high reliability point. Thus all coordinates of mixed blocks are defined easily. Then, all mixed blocks coordinates are defined, and these blocks are disposed at the required intermediate image.

Now equivalent blocks are found. Hence only the equivalent blocks are selected in the ascending order of likelihood. It is possible that some of the equivalent blocks overlay other blocks fully or partially, but the more likelihood the blocks remain visible. Operation with all possible blocks (candidates) guarantees that empty regions in an equivalent area are filled. The most important fact for rendering equivalent blocks is

to overlay the equivalent block on the mixed block. It is more accurate to say that the equivalent block overlays the part of the shadow block which includes the mixed block.

2. Results

Fig. 7 shows the image, "TOYS" shot with parallel axes of lenses. The intervals of cameras are about 35 mm, and three images are taken: left, right, and middle. By using the proposed algorithm from left and right images, the PSNR is 31 dB which is higher PSNR than the previous works in which the PSNR was 27.5 dB [11]. Table 1 summarizes the result of a few synthesized images with kinds of image.

IV. CONCLUSION

The visual estimation of synthesized intermediate images gives satisfactory results. Objective evaluations are carried out and more than 31 dB PSNR was obtained. From our point of view, the developed program is a good base for further development to achieve the possibility of an application to multi-view systems. It is necessary to test a developed program on a wide class of images with more different parameters. It is

TABLE 1. Experimental results.

source images	PSNR (dB)
"TOYS"	31
"DOLL & FLOWERS"	33
"FLOWERS"	32.3

possible that some complicated conditions occur during processing, so additional development will be required. The problem of processing motion images was not considered in this work. This problem is connected with hardware development, and the change of programs will be required.

V. ACKNOWLEDGEMENT

This work was supported by the Advanced Technology Development Program of the Ministry of Commerce, Industry and Energy under Contract No. UCM0941-6719-2.

*Corresponding author : eveduck@hanmail.net.

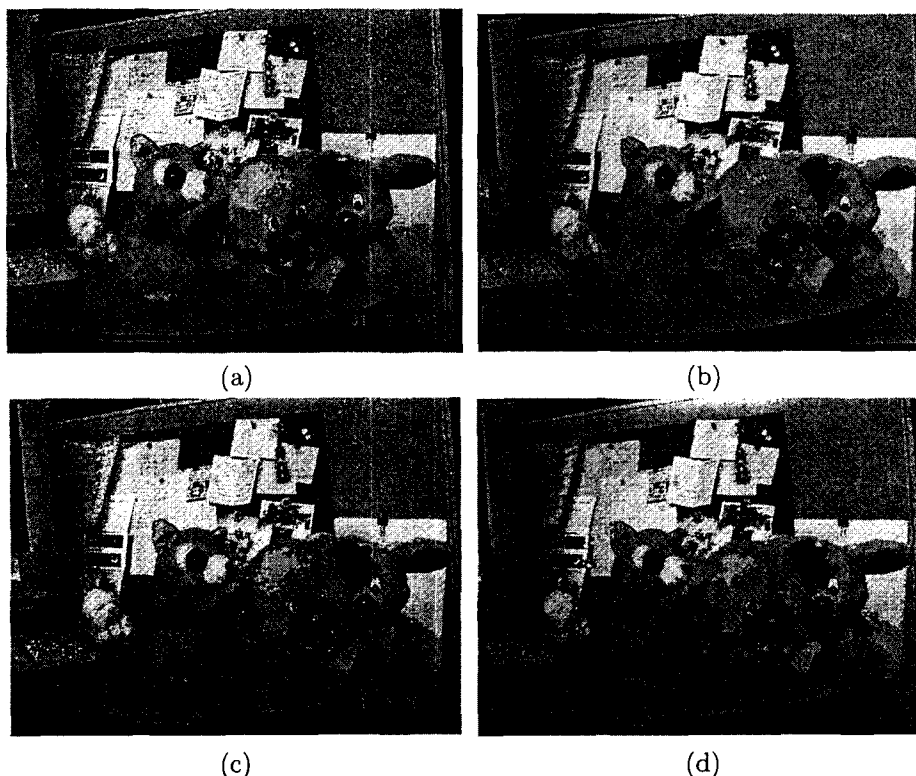


FIG. 7. (a) Original left image. (b) Original right image. (c) Original middle image. (d) Synthesized middle image.

REFERENCES

- [1] K. Susami, S. K. Kim, and J. Y. Park, *Journal of Optical Science & Technology*. **5**, 7 (2001).
- [2] K. T. Kim, *Journal of Optical Science and Technology*. **5**, 22 (2001).
- [3] K. T. Kim, V. Komar, and S. Ozerov, in *Proc. ODIP* (2002).
- [4] V. Komar and K. T. Kim, *Asia Display/International Display Workshop '01, Nagoya, Japan* (2001).
- [5] J. H. Kwak, K. T. Kim, and V. Komar, in *Proc. of the 9th Conf. on Optoelectronics and Optical Communications*, **9**, 133 (2002).
- [6] J. H. Kwak and K. T. Kim, *Journal of Hannam University* (2002).
- [7] J. H. Kwak and K. T. Kim, in *Proc. of Multimedia society of Korea* (2001).
- [8] K. T. Kim, *Final Report of Advanced Technology Development Program of the Ministry of Commerce, Industry and Energy* (2002).
- [9] T. Kobayashi, T. Fujii, T. Kimoto, and M. Tanimoto, in *Proc. of 3D image Conference '99* (1999).
- [10] J. S. Mcveigh, M. W. Siegel, and A. G. Jordan, *Image Communication* **9** (1996).
- [11] Z. Z. Kim, K. T. Kim, in *Proc. of Multimedia society of Korea* (2001).