

Novel Frame Interpolation Method for High Image Quality LCDs

Goh Itoh and Nao Mishima

Abstract

We developed a novel frame interpolation method to interpolate a frame between two successive original frames. Using this method, we are able to apply a double-rate driving method instead of an impulse driving method where a black frame is inserted between two successive original frames. The double-rate driving method enables amelioration of the motion blur of LCDs caused by the characteristics of human vision without reducing the luminosity of the whole screen. The image quality of the double-rate driving method was also found to be better than that of an impulse driving method using our motion picture simulator and an actual panel. Our initial model of our frame interpolation method consists of motion estimation with a maximum matching pixel count estimation function, an area segmentation technique, and motion compensation with variable segmentation threshold. Although salt and pepper noise remained in a portion of an object mainly due to inaccuracy of motion estimation, we verified the validity of our method and the possibility of improvement in hold-type motion blurring.

Keywords : frame interpolation method, double-rate driving method, motion blurm, motion picture simulator

1. Introduction

It is well known that the slow response time of liquid crystals (LC) causes motion blur of images on LCDs. We developed a driving method called LAO (Level Adaptive Overdrive) that reduces the LC response time to less than a single frame period (16.7 ms) [1]. On the other hand, Kurita et al. analyzed a hold-type display method in which the picture is sustained during the frame period and showed that this display method degrades motion pictures due to the characteristics of human vision. Although moving objects are displayed at the same position during a frame period in the hold-type display method, the human eye continues to move in order to trace them. A difference exists between the viewpoint and the moving object. Since shifted images that depend on the difference are displayed on the retina, hold-type motion blurring of images occurs (Fig. 1) [2]. Therefore, motion picture quality of LCDs is not better than that of CRTs whose driving method is of the impulse type, however fast response time of LCDs is.

Manuscript received May 29, 2004; accepted for publication July 17, 2004.
Corresponding Author : Goh Itoh
Multimedia Laboratories, Corporate Research & Development Center, Toshiba Corp. Kawasaki, Kanagawa, 212-8582, Japan.
E-mail : go.ito@toshiba.co.ip **Tel** : +81 44-549-2254
Fax : +81 44-520-1255

This motion blur can be reduced by an impulse driving method or a double-rate driving method (Fig. 2). In the impulse driving method, we insert a black frame between two successive original frames. On the other hand, in the double-rate driving method, we insert an interpolated

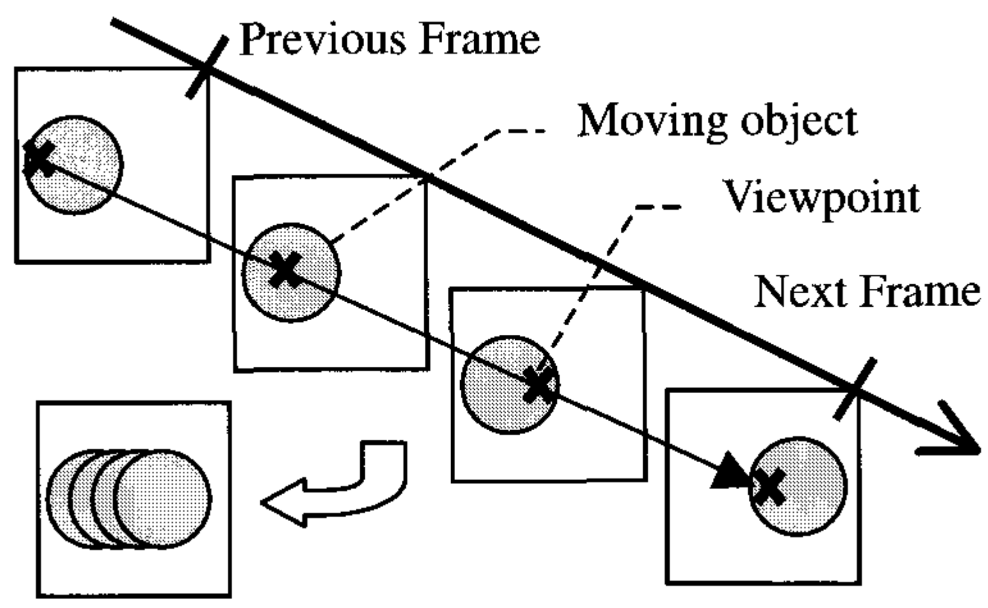


Fig. 1. Hold-type motion blurring.

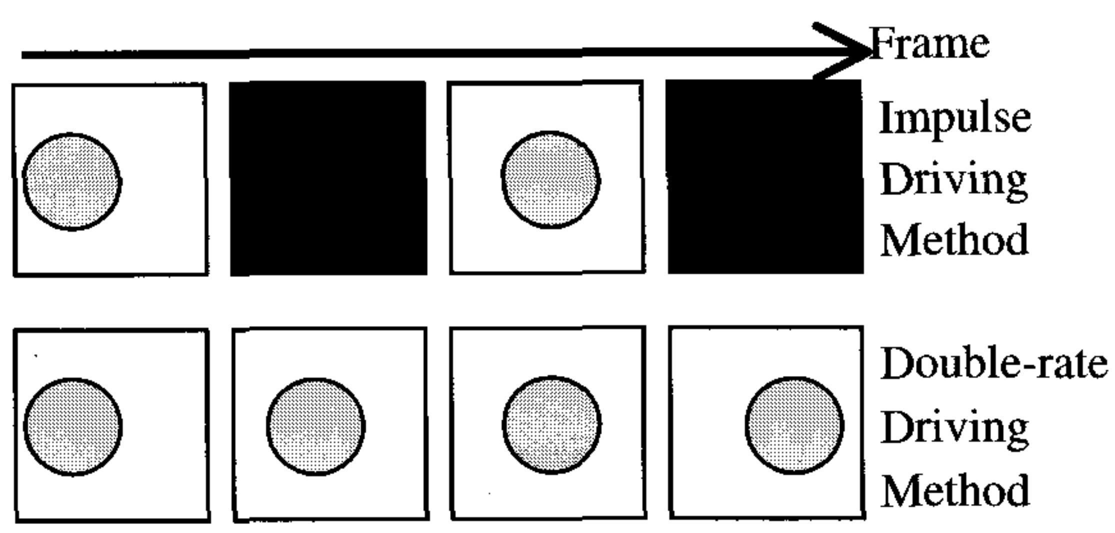


Fig. 2. Driving method for blurring.

frame between two successive original frames. In order to determine which driving method we should apply, 34 observers evaluated the simulated images subjectively with the universal motion picture simulator [3]. We decided to develop a frame interpolation method, judging from the evaluation results.

2. Subjective Motion Picture Evaluation

2.1 Evaluation with motion picture simulator

We have developed the universal motion picture simulator (UMPS) with an optical response measuring system. The UMPS has 3 key features as follows

[The UMPS features]

1) Measurement System for All Gray Response Data

2) Large Frame Memory (35s for HD picture)

3) High & Flexible Refresh Rate CRT (50~540 Hz)

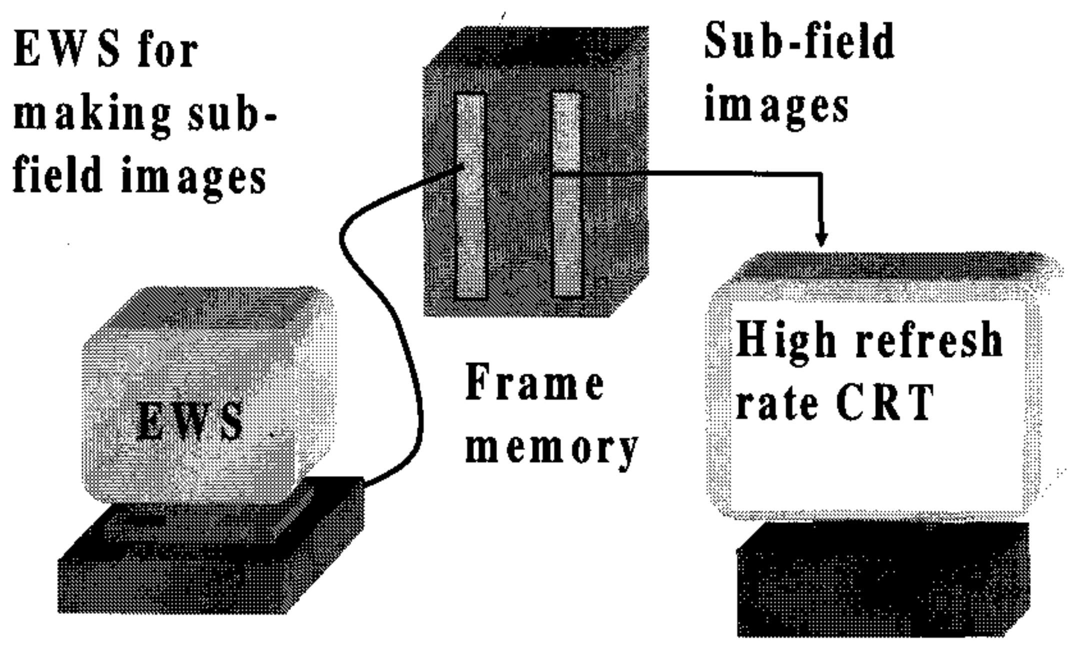


Fig. 3. Universal motion picture simulator.

Each step without the optical response measuring system as shown in Fig. 3 is explained briefly as follows. Referring to the response data table, sub-field images were made every 1/480 second with an EWS. Next, signals of sub-field images are stored in a large frame memory that is able to change refresh rate flexibly. Then, the images were displayed on a high refresh rate CRT monitor that changes refresh rate at 50 to 540 Hz. Finally we evaluate motion picture quality of images subjectively.

In this experiment on a motion blur, a still picture was scrolled horizontally as a motion picture. The simulation conditions were as follows:

[CONDITIONS]

1) DISPLAY SPECIFICATIONS

Vertical frequency : 480 Hz

Horizontal frequency : 130 kHz

Image size : 7 inch diagonal

Image resolution : 320 × 240

2) HOLDING TIME (the number of sub-fields during sustaining of the same picture) : 1, 2, 4, and 8 sub-fields

Defining the following equation as a holding rate, holding time of 1, 2, 4, and 8 sub-fields equals holding rate of 12.5(=1/8), 25(=2/8), 50(=4/8), and 100(=8/8) %H.R. respectively.

HOLDING RATIO (%H.R.)

=DISPLAYING PERIOD/(DISPLAYING PERIOD + NON-DISPLAYING PERIOD) * 100 (1)

3) LUMINOSITY

Maximum luminosity (white level) is the same for every display mode.

4) THE NUMBER OF FRAMES (fps)

For example, 240 fps means 240 picture frames per second.

5) OBJECTS MOVING SPEED : 16 degrees per second

We assumed that sport scenes have the objects moving speed of 10 to 20 degrees per second.

6) PICTURES : 2 kinds of CG picture & 2 kinds of natural picture

Fig. 4 shows results of average subjective evaluation value for 4 kinds of picture. Subjective evaluation value

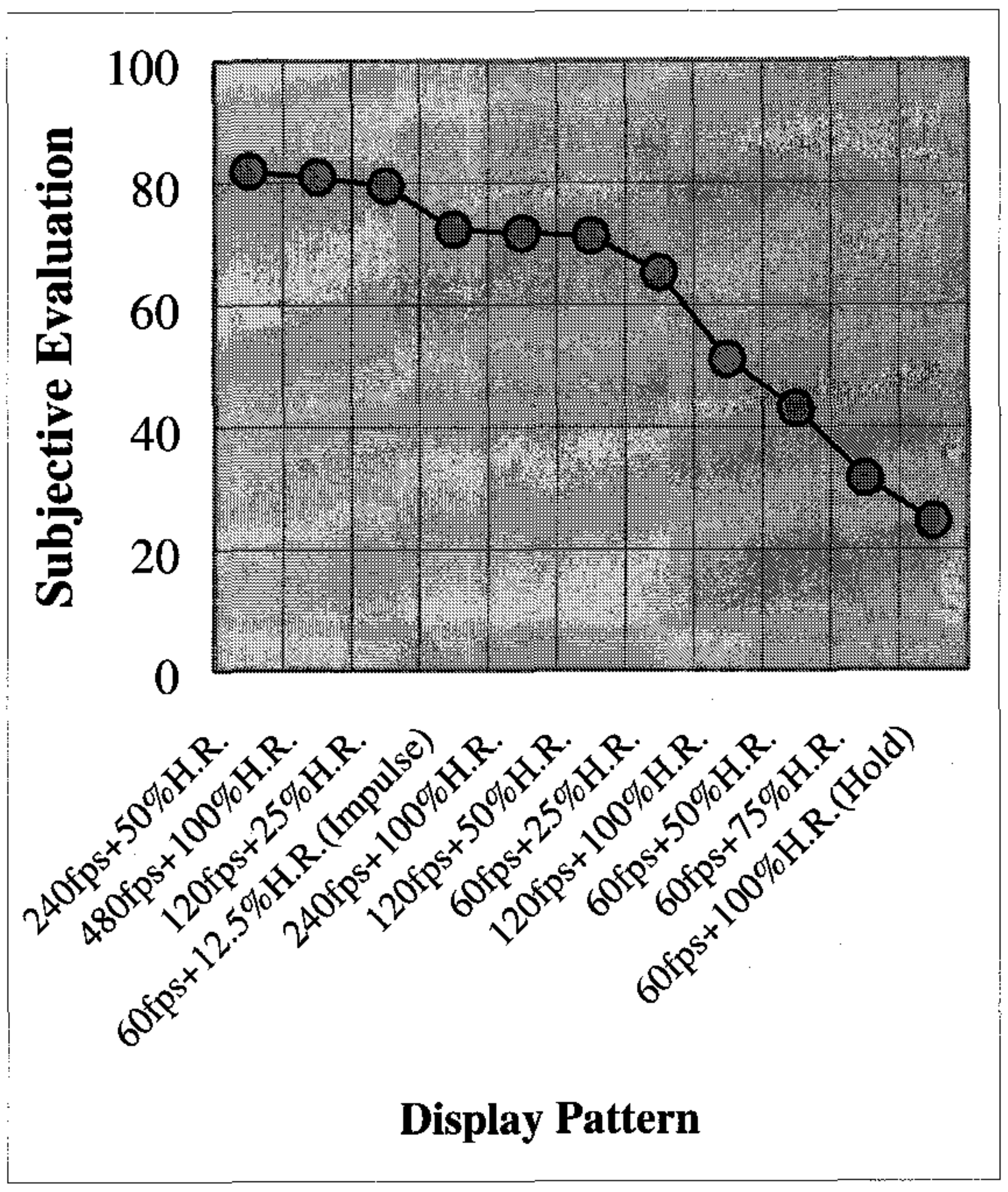


Fig. 4. Subjective evaluation for motion picture quality with the simulation.

depends on a motion blur and motion smoothness. Evaluation value "very good" means that a blur is almost equivalent to a still picture and objects move smoothly. For example, 240 fps+50 %H.R. and 120 fps+25 %H.R. are equivalent to 480 fps+100 %H.R. for motion picture image quality. In other words, they are the same in terms of hold-type motion blurring. Subjective evaluation value for 60 fps+12.5 %H.R. is, however, less than the above display mode, although a length of light emitting time for displaying one sub-field is the same, 2.08ms. This indicates that the more frames there are per second, the higher the motion picture image quality is. We also found similar characteristics of subjective evaluation for other display modes such as a difference between 120 fps+100 %H.R. and 60 fps+50 %H.R. Low frames per second leads to motion jerkiness caused by low correlation between a frame and the next frame. The motion jerkiness is likely to be remarkable for scrolling a clearer still image.

On the other hand, we confirmed that a hold-type motion blurring becomes large as holding ratio increases from 60 fps+12.5 %H.R. to 60 fps+100 %H.R.

2.2 Evaluation with actual LCD panel

Next, we investigated image quality with an actual LCD panel having high speed response property. The response time is about 16ms from white to black level. A black frame and a picture frame are displayed alternately at 60 fps+50 %H.R. Therefore, luminosity of a panel with 50 %H.R. is half that with 100 %H.R. Moreover, because LC response property is not ideal impulse-type, luminosity of 60 fps+50 %H.R. is reduced to about 30 % of 60 fps+100 %H.R. Fig. 5 shows comprehensive subjective evaluation obtained by comparing 120 fps+100 %H.R. or 60 fps+50 %H.R. with 60 fps+100 %H.R. Subjective evaluation value of 60 fps+50 %H.R. was lower than that of 60 fps+100 %H.R.

Judging from these results, we chose the double-rate driving method to realize 120 fps+100 %H.R.

3. Frame Interpolation Method

3.1 Motion estimation

In the double-rate driving method, we insert a middle frame between contiguous frames.

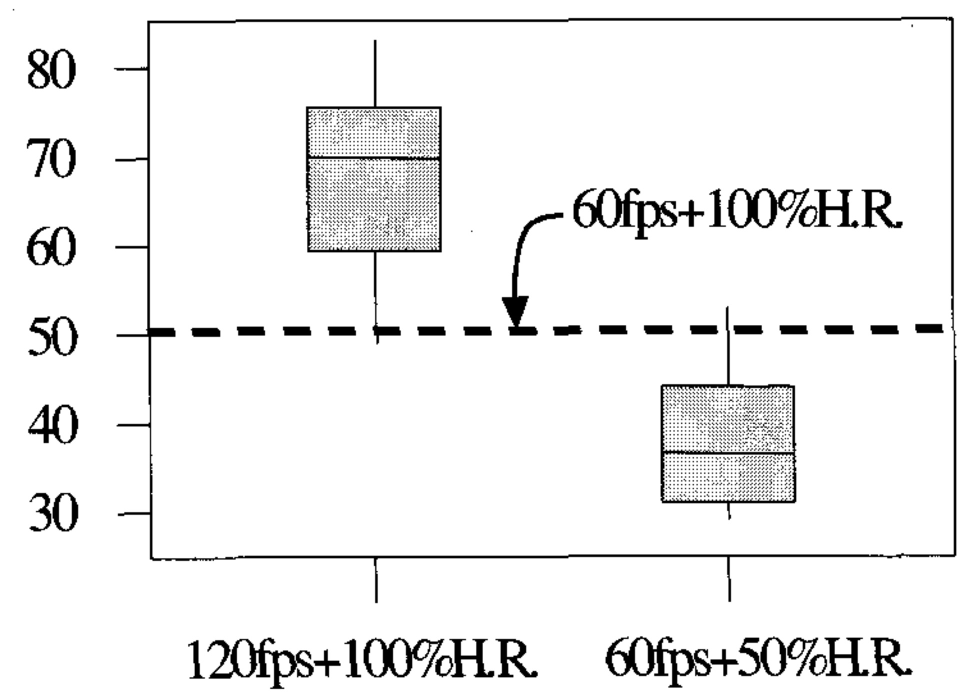


Fig. 5. Comprehensive subjective evaluation for display mode with an actual LCD panel.

The middle frame is created by a frame interpolation method consisting of motion estimation and motion compensation. We constructed frame interpolation algorithms based on the block-matching algorithm used for MPEG. In the block-matching of MPEG, high correlation blocks are searched by motion estimation. There are various estimation functions [4], sum of absolute difference estimation function $E_{SAD}(d)$ is usually applied to MPEG, because the focus has been on reducing bit rates.

$x \in R^2$ represents the coordinate in frame $t \in R$ and $s(x,t)$ represents the intensity value at $(x,t)^T$.

$$E_{SAD}(d) = \sum_{x \in B} |s(x-d,t) - s(s+d,t+1)| \; \mathrm{L} \qquad (2)$$

$$d = \arg \min_{d \in W} E_{SAD}(d) \; \mathrm{L} \qquad (3)$$

$$d \in R^2$$

represents a motion vector, B is a block and W is a search window.

Accurate motion estimation is a very important factor for motion compensation of the frame interpolation method. However, for example, different interpolated frames A and B with motion vectors V01 and V02, respectively, are able to be reconstructed as shown in Fig. 6. There is block distortion caused by low correlation block B1 in the interpolated frame A. Therefore, we needed another estimation function to focus on image texture. Maximum matching pixel count estimation function $E_{MPC}(d)$ depicted in Equation (4) to (6) is as follows.

$$E_{MPC}(d)=\sum_{x\in B}\sigma(x,d) \qquad (4)$$

$$\sigma(x,d)=\begin{cases}1 & |s(x-d,t)-s(x+d,t+1)|\le T1\\ 0 & otherwise\end{cases} \qquad (5)$$

$$d=\arg\max_{d\in W}E_{MPC}(d) \qquad (6)$$

Fig. 7 shows the example of block candidates B0, B1 and B2 in Fig. 6, and results of $E_{SAD}(d)$ and $E_{MPC}(d)$ (Searching Threshold: *T1* in equation (5) = 2).

Although $E_{SAD}(d)$ between B0 and B1 is smaller than $E_{SAD}(d)$ between B0 and B2, B1 is a raster texture. On the other hand, although average intensity value of B2 became higher than that of B0, $E_{MPC}(d)$ between B0 and B2 is larger than $E_{MPC}(d)$ between B0 and B1. $E_{MPC}(d)$ is likely to be more dependent on image texture than $E_{SAD}(d)$.

3.2 Area segmentation technique

In addition to the above matching characteristics, to avoid block distortion caused by inaccurate motion compensation for boundaries of moving objects when two or more moving objects exist in a block, we applied an area segmentation technique to the basic block-matching algorithms. As shown in Fig. 8, the first block matching between two contiguous frames was performed geometrically and symmetrically centered on an interpolated frame.

Secondly, we divided a block into a high correlation area (dotted area) and a low correlation area (black area) depending on a segmentation threshold *T2* which can be the same value as *T1* in equation (5).

Thirdly, in the second block matching, high correlation blocks (hatched area) whose shape is the same as the low correlation area are searched between previous frame and next frame using the same searching algorithm as for the first block matching. Next, the first motion vector detected by the first block matching was assigned to the dotted area and the second motion vector detected by the second block matching was assigned to the hatched area.

3.3 Motion compensation

Image data of pixels that motion vectors pointed to was pasted to an area in the interpolated frame according to each motion vector. The great difference between the interpolation method and MPEG is with or without residuals used in MPEG. There are no residuals for the interpolation method, because an original image is not available at an interpolated frame.

Finally, an area in the interpolated frame that image data was not pasted to was allotted from peripheral pixels or contiguous frame using a pixel interpolation method as follows.

Fig. 9 shows pixel interpolation for the area without image data obtained by first and second block matching. Gray pixel in the interpolated frame is calculated by non-linear averaging such as median filter or linear averaging. Image data, using non-linear averaging, are retrieved from peripheral pixels (hatched pixels) with image data obtained by first and second block matching and contiguous pixels

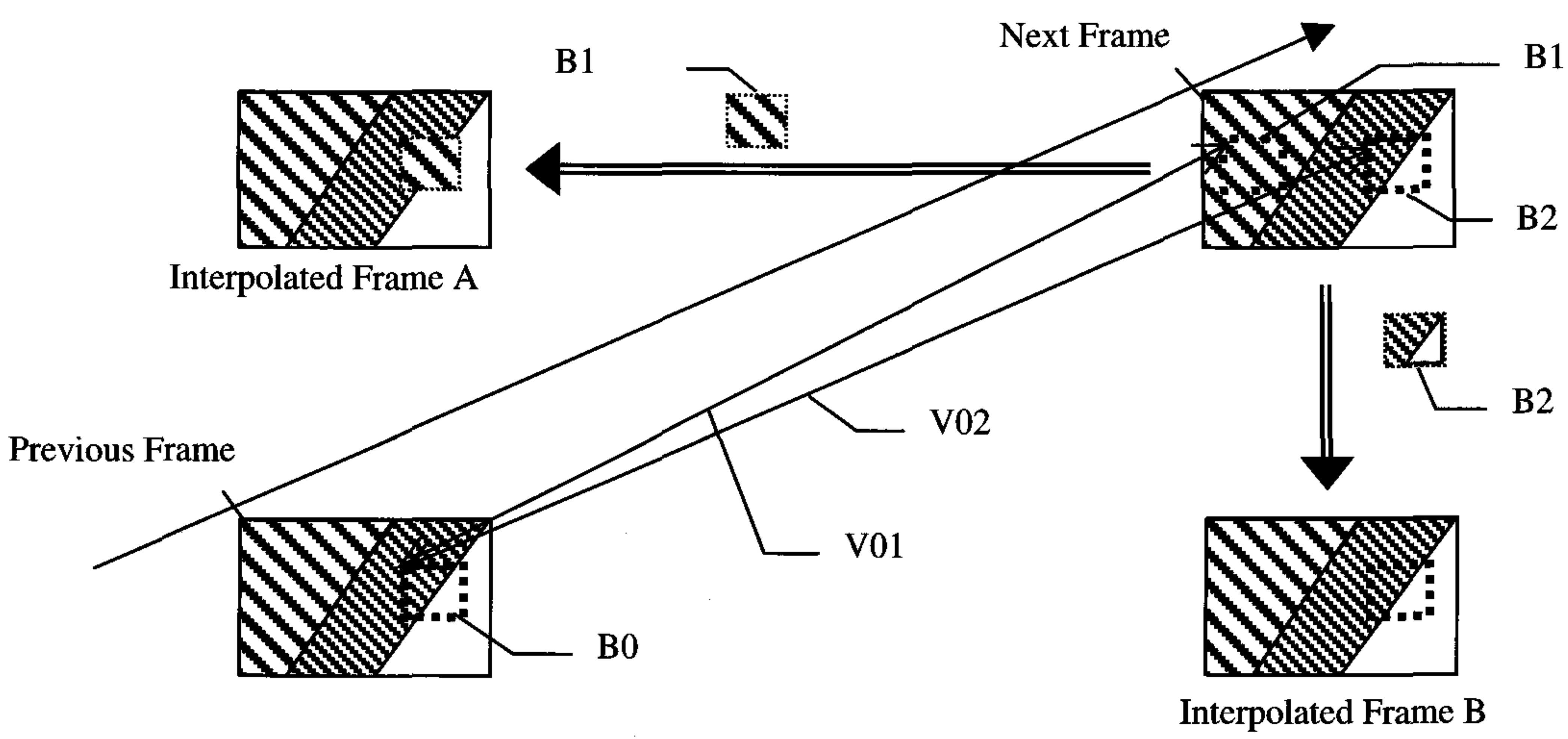


Fig. 6. Block distortion caused by inaccurate motion compensation for boundaries of objects.

(dotted pixels) which are pointed to by first and second motion vectors on original frames. If there is no image data in a peripheral pixel, we ignore the peripheral pixel.

There are several pixel interpolation methods. We intend to investigate another method based on motion objects segmentation.

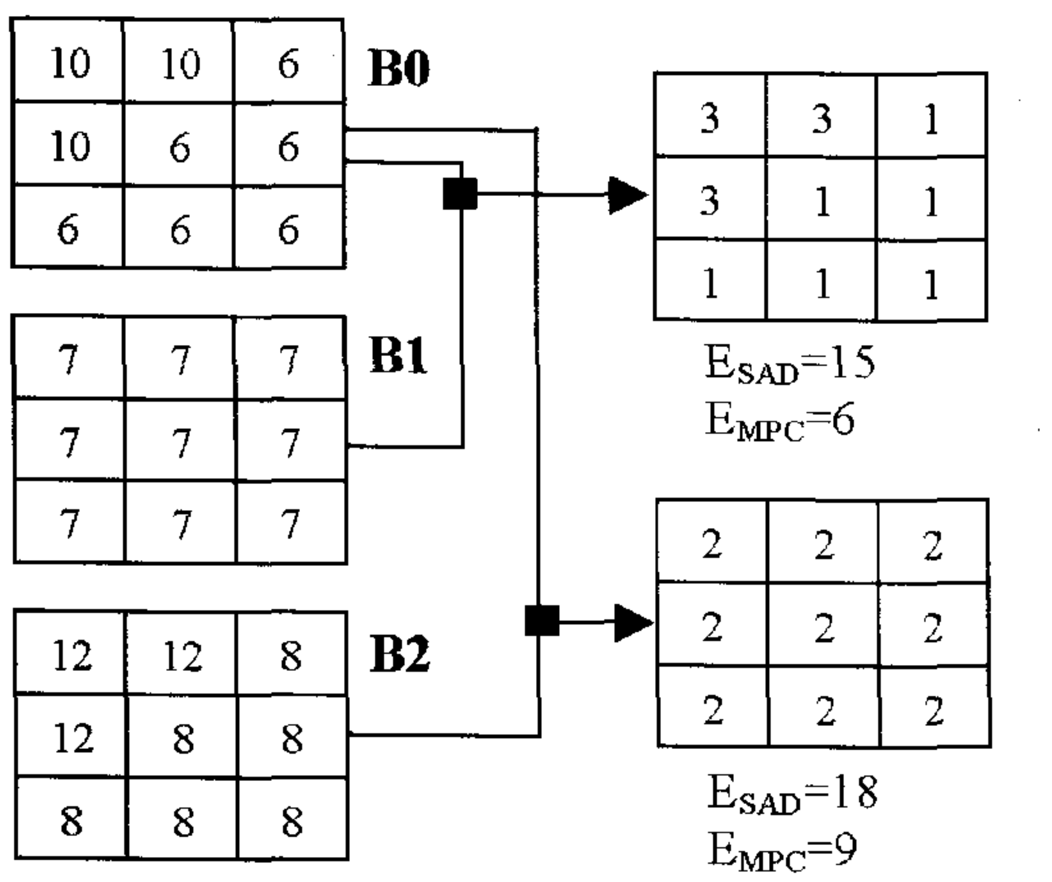


Fig. 7. Correlation between block candidates obtained by calculating $E_{SAD}(d)$ or $E_{MPC}(d)$.

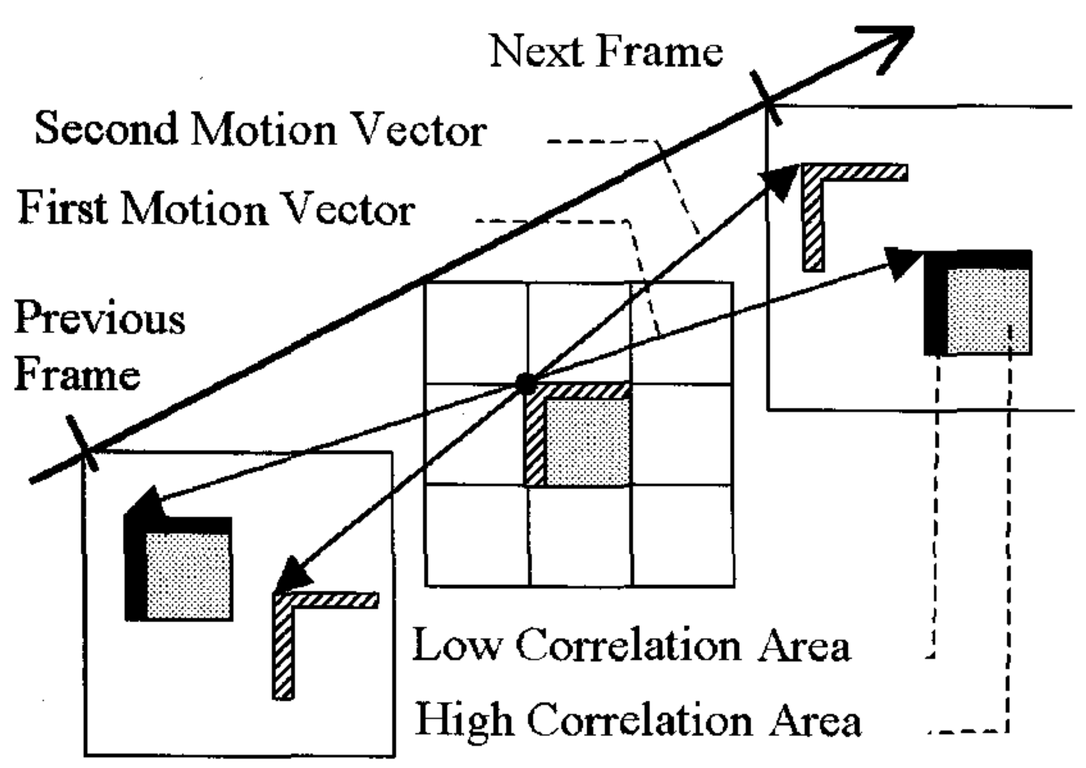


Fig. 8. Proposed interpolation method.

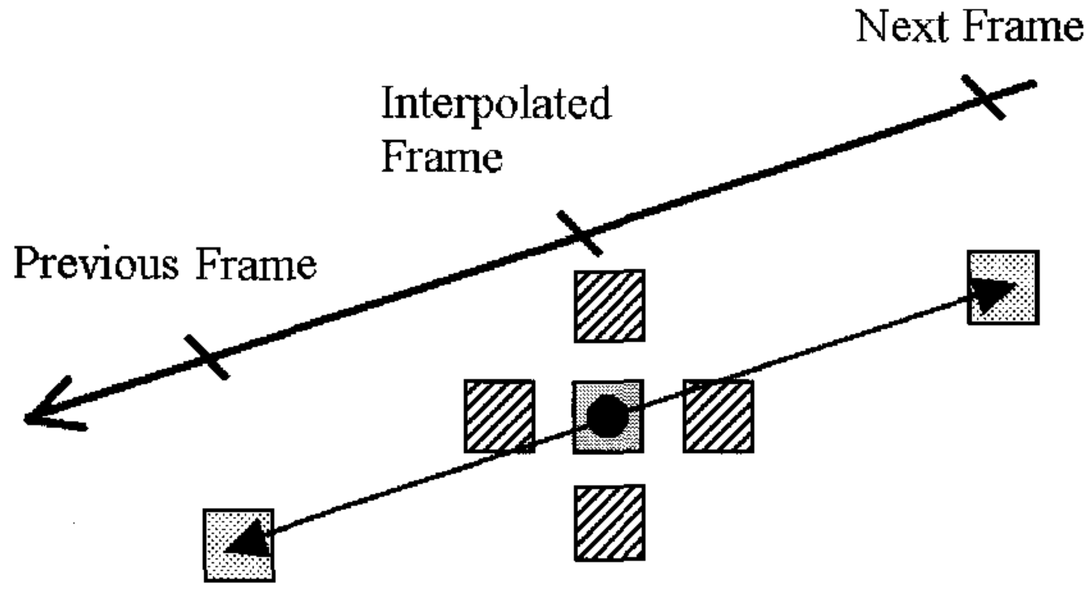


Fig. 9. Pixel interpolation.

3.4 Variable segmentation threshold

We examine the thresholds *T1* and *T2*. *T1* is the searching threshold for motion estimation and *T2* is the segmentation threshold for motion compensation. It is not necessary for *T1* to be equal to *T2*. If *T1* is large, block matching does not depend on texture and if *T1* is small, block matching is sensitive to slight intensity change such as noise. On the other hand, if *T2* is large, motion compensation is insensitive to texture and noise and if *T2* is small, the number of pixels without image data becomes large. Too small *T1* and too small *T2* increase pixels without image data, which leads to salt and pepper noise due to low correlation between a pixel with interpolated image data and a peripheral pixel with obtained image data. For example, if the object deforms slightly, the edge of the object shifts around a few pixels. B0 in the previous frame and B2 in the next frame are retrieved and then the absolute difference value per pixel is calculated. Two pixels (dotted pixels) in the center of a block become pixels without image data as shown in Fig. 10. These pixels usually appear on the edge of objects, because the luminosity difference between neighboring pixels in the edge region is generally higher than that in the raster region. Therefore, we applied a different segmentation threshold between the edge region and the raster region. The threshold *T2* for the edge region is higher than that for the raster region.

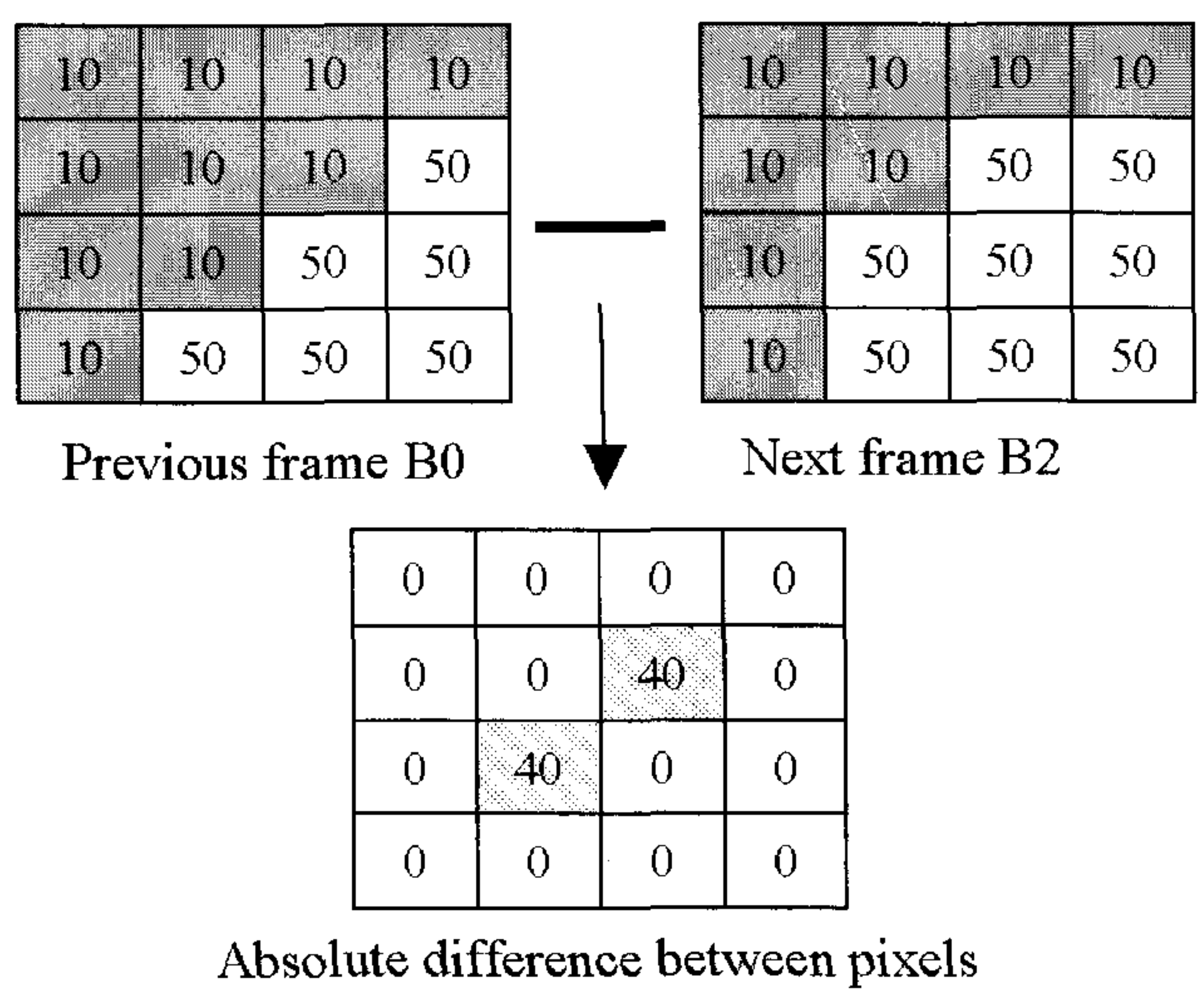
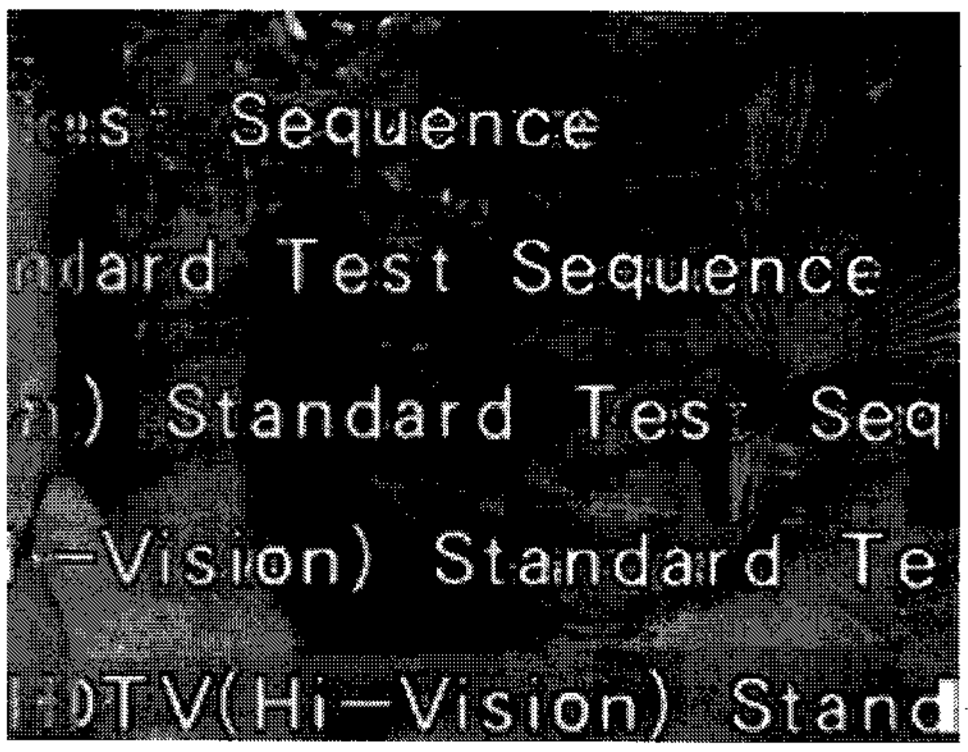
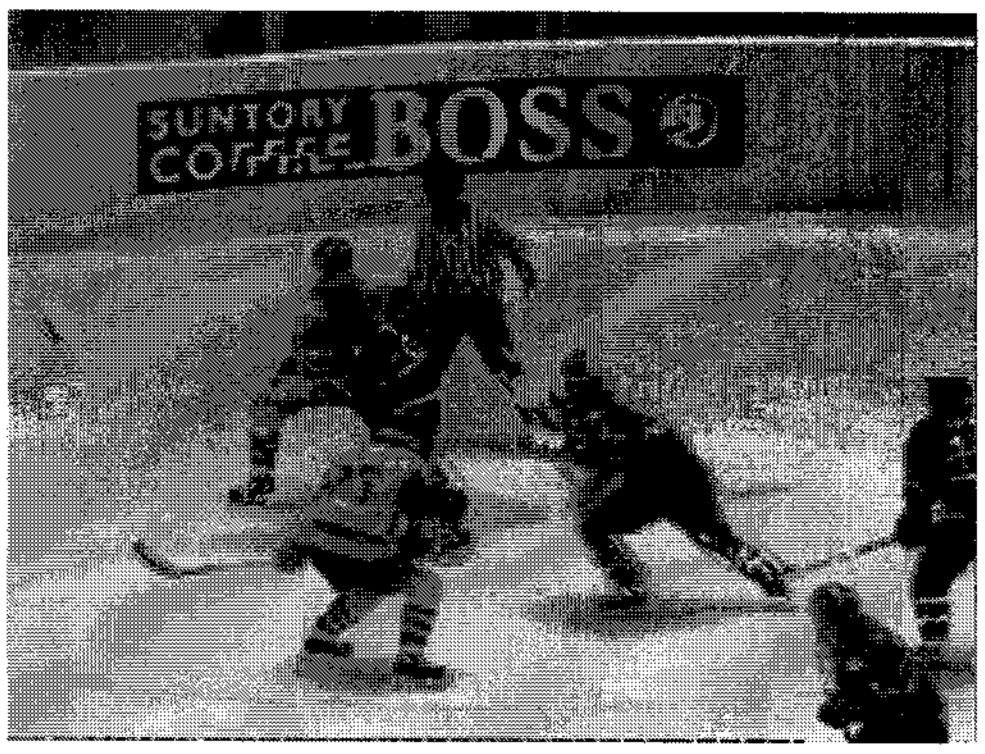


Fig. 10. Salt and pepper noise.



(a)



(b)

Fig. 11. Interpolated pictures: (a) Bronze; characters move right to left on a background image with natural plants. (b) Hockey; Players move in various directions.

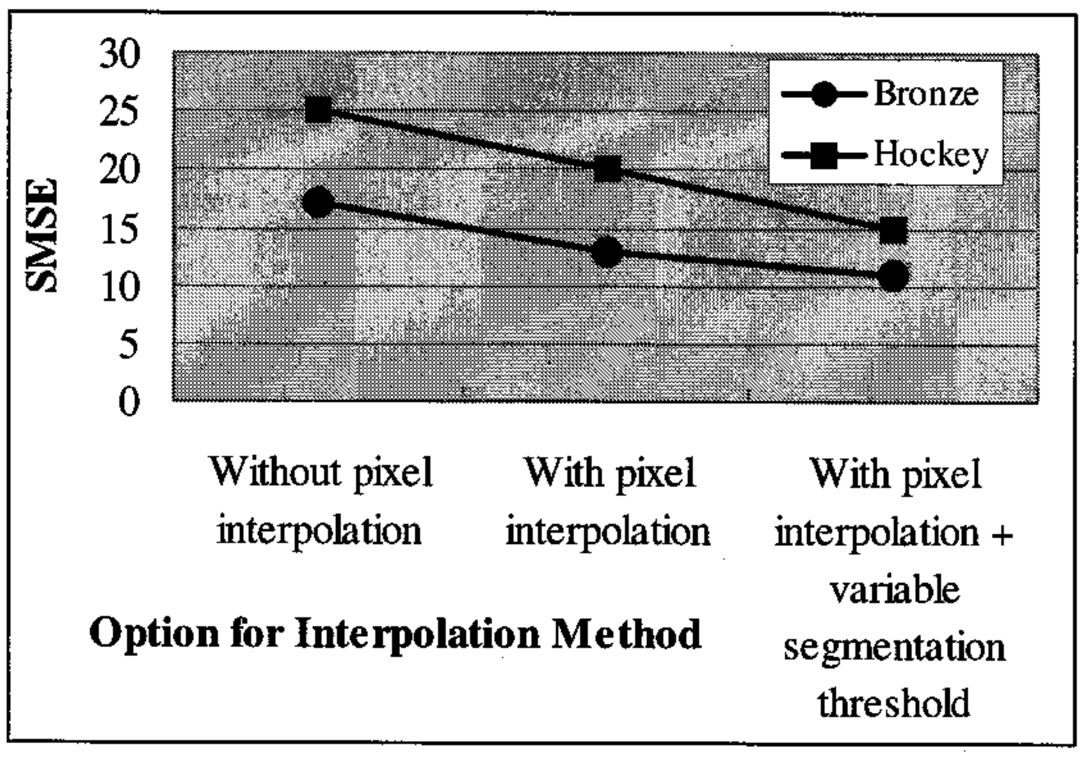


Fig. 12. SMSE of interpolated pictures.

4. Results

Fig. 11 shows an interpolated picture realized by using our interpolation method. Salt and pepper noise remained between characters and a portion of a character such as "Test" on the right side disappears in the "Bronze" picture. Similarly, a foot of an umpire or a player disappears in the "Hockey" picture. These image degradations were caused by segmentation threshold and occlusion. Occlusion is a problem that we cannot solve with 2 frames. Moreover, "COFFEE" on the left side has been broken. This image degradation mainly results from inaccuracy of the first motion estimation.

We have used image quality measures: Subjective Mean Squared Error (SMSE) [5]. The SMSE which we used is defined by:

$$SMSE = \left(\frac{1}{MN} \sum_{x=1}^{M} \sum_{y=1}^{N} \left| g(x,y) \right|^3 \right)^{1/3} \quad \cdots\cdots\cdots\cdots\cdots (7)$$

$$g(x,y) = M_{2\times2}(s_{ref}(x,y)) - M_{2\times2}(s_{test}(x,y)) \quad \cdots (8)$$

$M_{2x2}(s(x,y))$ denotes a linear average in a 2×2 pixel window to describe the low-pass filter of the eye. Image quality has been improved by adding options to the interpolation method as shown in Fig. 12. Here, we set image data of pixels to which no motion vector is allotted to 0 for Without Pixel Interpolation. However, more improvement is necessary for reducing SMSE to less than 10, which is our initial target.

5. Conclusion

We found that the double-rate driving method is better than the impulse driving method, judging from motion picture simulation.

Therefore, we developed an initial model of the frame interpolation method between two successive original frames. And we verified the validity of our method and the possibility of improvement in hold-type motion blurring.

Salt and pepper noise remained in characters and a portion of an object mainly due to inaccuracy of motion estimation. Our next task is to find a solution for the problem of image degradation. The improved and robust method leads to realization of the double-rate driving method for ameliorating motion blur for LCDs.

References

[1] H. Okumura and H. Fujiwara, in *SID Digest* (1993), p. 335.
[2] H. Ishiguro and T. Kurita, *Tech. Rep. of IEICE, EID96-4*, 19 (1996).
[3] G. Itoh, M. Baba, and H. Okumura, in *Asia Display/ IDW'01* (2001), p. 1633.
[4] A. M. Tekalp, "Digital video processing", Prentice Hall, Chapter 6 (1995).
[5] H. Marmolin, IEEE Trans. on System, Man and Cybernetics, **16**, 486 (1986).