

Music Score Recognition Based on a Collaborative Model

Van Khien Pham and GueeSang Lee*

*Chonnam National University, Kwangju, Chonnam National University, Kwangju,
South Korea, South Korea
poldpham@gmail.com, gslee@chonnam.ac.kr*

Abstract

Recognition musical symbols are very important in music score system and they depend on these methods of researchers. Most of existing approaches for OMR (optical music recognition) removes staff lines before symbols are detected, therefore the symbols can get damaged easily. Another method recognizes symbols without staff line removal but all of them have a low accuracy rate and high processing time for recognizing symbols. In this paper, none staff removal and staff removal are suggested and these new methods are proposed to improve appreciation result of symbols. A lot of symbols are detected before deleted staff line as vertical lines, note head, pitch, beam, tail and then these staff lines are removed to identify other symbols using connected component. The proposed method is applied to the Samsung smart phone which embeds a high resolution camera. Experimental results show that the recognition rate is higher than existing methods and the computation time is reduced significantly.

Keywords: *Line Detection, Template Matching, Music Scores, Staff line Detection, Note Head, Projection, Staff Line Removal*

1. Introduction

For the purpose of education and entertainment, OMR systems have amount of interest for many researchers. These systems are integrated in smartphone, which allow playing many music types by capturing a music sheet. Another way, OMR system is useful to compress and store from picture to xml file.

There are many existing papers recognizing a music score but all methods are implemented from a printed image [2-4] or hand written image [5]. That means music symbols in the input images are very clear. They are not affected by illumination and distortion. Other methods mentioned problem with non-linear distortion images captured by digital cameras. However, the symbol recognition is applied after staff lines removed [6-7,10]. Thus some symbols are damaged and the recognizing result is not correct. Besides that, another method does not apply the staff line removal [1] but it has many problems in identifying symbols. For instance, stem line detection uses vertical line projection and counting black pixel numbers. Thus, it leads up to many noises, stem lines are detected incorrectly.

Stem line detection is the most important step in the music score recognition. If stem lines are detected wrong, note head cannot be recognized. Note head detection uses template matching method for scanning template around stem line. It is sometimes incorrect and calculating time is so slow. Tail symbols are determined after eliminating the horizontal lines around tails. Therefore symbols are destroyed and detected incorrectly. All of existing methods were successful, but none of them has provided nearly close to a high precision.

In this paper, we give some methods to increase the performance of recognizing symbols in the images captured by the camera of the mobile phones. Based on this study, stem and bar line detection are given the longest line for scanning vertical projection. It is

easy to remove stream lines such as clef and rest symbols. Using template matching to recognize note head in the top or bottom position. Whole note usually stays between the consecutive bar lines. To detect whole note, we count the number of white pixels inside note when scanning horizontal projection. We use RLE method to detect beam and tail symbols without horizontal line removal. All symbols identified above are recognized without staff line removal and then staff lines will verify other symbols.

The processing speed is significantly increased in this study. It is really necessary for OMR system to run on mobile phones with the limited processing speed and memory. Our OMR system can handle many types of music sheet (for example: normal, curved, perspective). This paper is described in Figure 1 and is organized as follows. Section 2 presents the proposed method. Section 3 shows the result of experiment. Finally, Section 4 gives the conclusion.

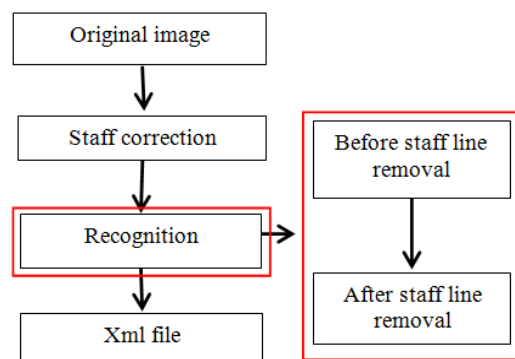


Figure 1. The Process of Our System

2. Proposed Method

2.1 Symbol Recognition without Staff Line Removal

2.1.1 Vertical Line Detection: Firstly, we use horizontal projection to detect staff line. When staff line is detected, vertical lines in each stave are determined by using vertical projection. A stave is the five closest consecutive staff lines. Firstly, left space, right space and search range need to be verified as shown in Figure 2.

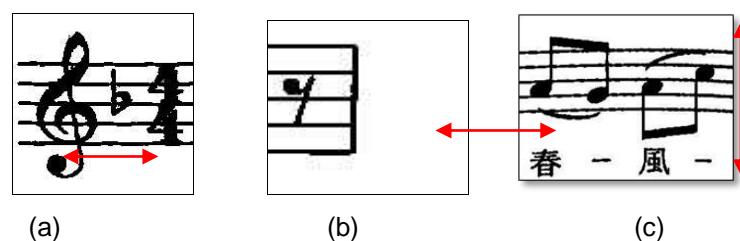


Figure 2. Left Space, Right Space, Search Range Information

Left space means the expected distance from the beginning of a staff line to the right hand side of the clef. Left space is used to skip the histogram computation for the clef. Right space is a distance from the end staff line to the right border of music sheet. Search range means the expected distance from the top staff line to bottom staff line of each stave. It is used to scan vertical projection in this range.

Each vertical line is taken the longest connected line to remove noise. The musical scores are classified among rest, sharp symbols and stem, bar lines based on vertical histogram Figure 3. In other words, stem and bar lines show a high histogram.

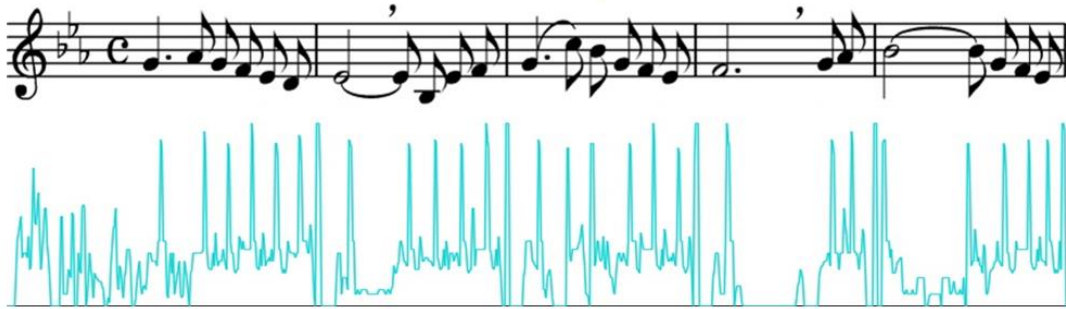


Figure 3. Histogram of Vertical Projection

The values appropriate when they satisfy the following equation:

$$\tau = urH + vrS$$

The threshold τ for vertical lines is based on the height of staff line and the distance between staff lines. where rH , rS are the thick staff lines and distance of two consecutive staff lines respectively as Figure 4 illustrates. u , v are defined values.

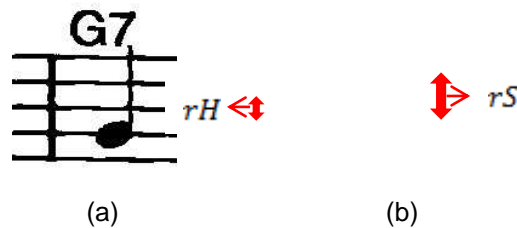


Figure 4. Staff Height and Staff Space Information

Algorithm for vertical line detection

- Step 1: Determine staff line information
- Step 2: Use vertical histogram for each stave
- Step 3: Use local maximum to choose the longest line
- Step 4: Give a condition about height for vertical line

In this paper, α , β are selected as 3 and 2 respectively in practice to ensure that all vertical lines are collected precisely. Then, stem lines and bar lines are needed to distinguish. Bar lines can be recognized easily by specific features. These lines have two peaks located on the top and bottom staff lines of each stave. Beside left and right hand side of bar lines do not have too many black pixels. Therefore, stem lines are collected precisely. A musical note is classified according to stem line position.

2.1.2 Musical Notes Detection: Musical notes have black notes, white notes, multi black notes and multi white notes.

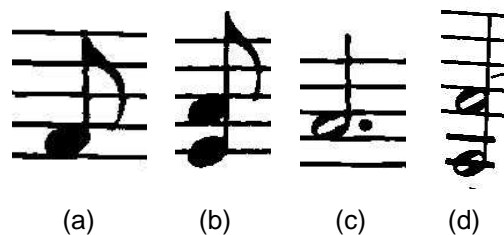


Figure 5. Black Note, Multi Black Note, White Note and Multi White Note

Black and white notes Figure 5 (a),(c) are usually located on the bottom left hand side or the top right hand side of the stem lines. Template matching for black note head is applied for each position of vertical line. The template is put at two positions bottom left hand side and top right hand side. If the matching result is higher than the threshold δ , black notes are detected. In this case, threshold δ is determined equal 0.75. The algorithm for musical notes recognition is the same following:

Algorithm for musical note recognition

Step 1: Create a template

Step 2: Black note detection

Use template matching method

Step 3: White note detection

Fill in the hole

Use template matching method

Step 4: If black notes are detected then

Detect multi black note using template matching

Step 5: If white notes are detected then

Detect multi white note using template matching

For white note head, after finding out all black note head, the rest of stem lines is checked for holes around them and matched becoming black note head after filling holes [9]. When black and white notes are detected, multi black and white notes (Figure 5 (b), (d)) are determined if another note is found around this stem using template matching method.

2.1.3 Whole Note and Dot Detection: To detect whole note, we base on the features of whole note in the music score. If whole note exists, it picks up entire measure Figure 6(a). Therefore, after determining black and white note, the rest of vertical lines are bar lines. If there is not any black or white note between two bar lines, the whole note can appear. To detect the position of whole note, we extract the measures which do not have black or white note. The position of whole note is detected by determining holes based on scanning horizontal projection and counting number of white pixels in the hole of Figure 6(b). The algorithm for whole note detection is the following:

Algorithm for whole note detection

Step 1: Determine bar lines information

Step 2: If there is no stem line between two consecutive bar lines then

Scan horizontal projection

Count number of white pixels

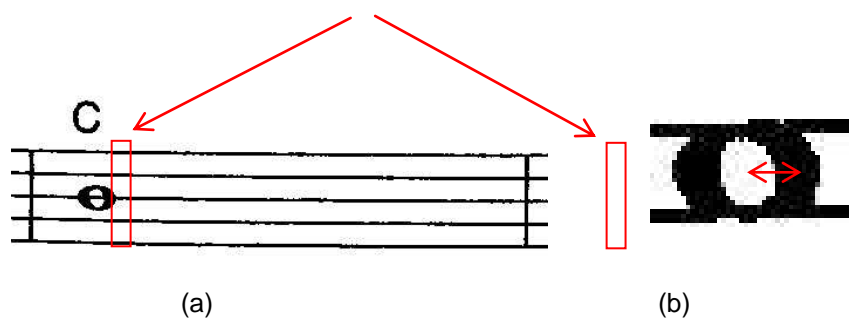


Figure 6. The Measure of Music Score with a Whole Note and the Hole Position of this Note (a) and Whole Note (b)



Figure 7. The Dot's Position

To recognize dot symbol, this symbol has characteristic about the position. According to the music score's rule, the dots are always located directly behind and middle of the note head (both black and white note) Figure 7. Therefore, based on the information about the position of note head determined above, the interesting region around behind the head note is verified. RLE is used to find out the black regions with the size of height and width approximating rS .

2.1.4 Pitch Detection: The main target of music symbol recognition is that we need to determine which type of symbol belongs to and which position a symbol is located. Therefore, pitch detection after recognition is an important step in a music symbol recognition system.

To determine the pitch, we base on the position of head note recognized in the above step. After getting the center position of note head, RLE is applied to find out the above and bottom position of note head. Then center position of note head is calculated again by taking average of above and bottom boundary according to vertical line Figure 8(a). A mention map that covers all possible positions of note head is created. The lowest pitch is -4 corresponding to lowest position and the highest pitch is 14 corresponding to the highest position of each staff. Each pitch level has a stage which equals to a half of the distance between consecutive staff lines. Each note has a pitch corresponding to the position of staff line being nearest to the center of note head Figure 8(b).

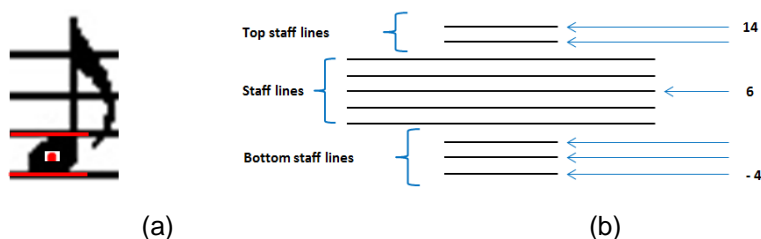


Figure 8. The Above and Bottom Boundary of Note Head and the Pitch of Note Center (a) and the Pitch Map for Pitch Detection (b)

2.1.5 Beam and Tail Detection: Other components of major symbols in a music score include tags and beam. In this Section, the information of detected notes is already verified. Beam notes have two types: single beam note is combined between two consecutive black notes and multi beam note is combined between two consecutive multi black notes. A beam is a thick line crossed to two stem lines which have two consecutive notes located at the bottom or top together as Figure 9. To detect beam, we use RLE for between two stem lines to find out the beams with the thickness of beams being larger than two times of the staff height (rH).

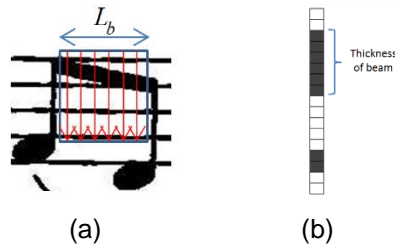


Figure 9. Beam Note (a) and RLE (b)

If the number of black pixels runs equals the beam length (L_b), beam line will be detected.

To detect tails Figure 10 and recognize how many tails are assigned to one note symbol, we firstly extract the part can have tails based on the center position of note head. Tails are usually located a top or bottom stem line on the right side Figure 10. We use RLE for right side of stem to find out the tails with the thickness of tails being larger than two times of the staff height (rH). It is important that staff lines are not removed in the process detection because tails are damaged and separated into many parts.

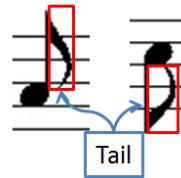


Figure 10. The Tails of Note Symbol in Music Score

Algorithm for beam and tail detection

Step 1: Determine note head position

Step 2: If two consecutive note heads are the same bottom or top position

Using RLE for detecting beam

Step 3: If these note heads do not have crossing thick line then

Detecting tail using RLE for right stem side



Figure 11. Clef (a), Sharp(b), Flat(c) and Natural Symbols(d)

Besides, Clef is always located at the beginning of each staff line, and its height is always higher than the height of staff line. For recognizing other symbols such as sharp, flat, natural symbols Figure 11 (b, c, d), we use vertical histogram and based on the height, width and characteristic of a symbol.

2.2 Symbol Recognition with Staff Line Removal

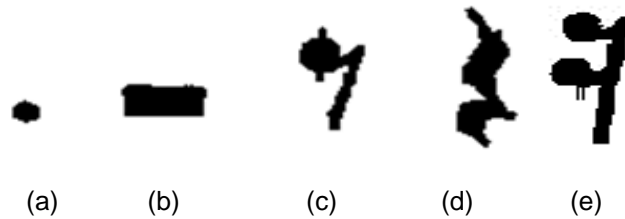


Figure 12. Dot (a), Rest2 (b), Rest4 (c), Rest8 (d) and Rest16 Symbols (e)

After these symbols are detected without staff line removal, staff removal is done with stable path and line tracking. Then, ROI is labeled by taking the advantage of typesetting rules of musical notations as follows : the top-left of the rectangle should be as high as size of full staff size, which is four staff space plus five staff line thickness; the bottom left is twice as high as the height of two consecutive staff lines. Connected component is analyzed to draw bounding boxes for other symbols and we have some information about height and width these symbols to recognize them Figure 12.

$$+ \text{ For Dot: } \quad H_D < 0.75 * rS ; \quad H_D > 0.25 * rS \\ W_D < 0.75 * rS ; \quad W_D > 0.25 * rS$$

Where H_D and W_D are the height and width of Dot symbol.

$$+ \text{ For Rest2: } \quad H_{R2} < 0.8 * rS ; \quad H_{R2} > 0.3 * rS \\ W_{R2} < 1.25 * rS ; \quad W_{R2} > 0.75 * rS$$

Where H_{R2} and W_{R2} are the height and width of Rest2 symbol.

$$+ \text{ For Rest4: } \quad H_{R4} < 2.3 * rS ; \quad H_{R4} > 1.75 * rS \\ W_{R4} < 1.5 * rS ; \quad W_{R4} > 0.5 * rS$$

Where H_{R4} and W_{R4} are the height and width of Rest4 symbol.

$$+ \text{ For Rest8: } \quad H_{R8} < 3 * rS ; \quad H_{R8} > 2.3 * rS \\ W_{R8} < 1.4 * rS ; \quad W_{R8} > 0.5 * rS$$

Where H_{R8} and W_{R8} are the height and width of Rest8 symbol.

$$+ \text{ For Rest16: } \quad H_{R16} < 3 * rS ; \quad H_{R16} > 2.3 * rS \\ W_{R16} < 2.2 * rS ; \quad W_{R16} > 1.4 * rS$$

Where H_{R16} and W_{R16} are the height and width of Rest16 symbol.

3. Experimental Results

3.1 Dataset

We captured 470 images in middle school and high school of Korean children songs using Galaxy Note 3 mobile phone with various morphologies and output image is displayed as in Figure 13. They have the different size of shape, illumination, distort and view point, *etc.* The user interfaces for capturing the image and displays the recognition result is implemented on Java. We use android NDK to support in C++ and support the fastest recognizing time.

+ Sample 1(Moon river)

Moon river
H. 텐서나 작곡 | 윤경미 역사

이 노래를 리코디나 피레, 오카리나 등의 기타 악기로 연주해 보자.
한구들과 함께 영화에 사용된 음악 중 인상 깊었던 곡을 시로 이야기해 보고, 그 음악이 나오는 부분을 동영상 작곡으로 다시해 보자.

수상하였으며, 지금까지 널리 불리고 있다.

이 노래는 1961년 데이비드 멀헤 주제가가 작곡
수상하였으며, 지금까지 널리 불리고 있다.

문 리 버 제 강을 건너 나 그 데 에 게 가 리
Moon Riv - er, wid - er than a mile. I'm cross - in' you in style some
라 - 내 꿈 그 데 만 읊 사 영 하
day - Oh, dream mak - er, you heart break - er, wher -
리 언 제 나 - 그 데 와 - 함께 이 세 상
ev - er you're go - in' - I'm go - in' your way. Two drift - ers,
F C F C Dms Er Am
방 광 하 여 도 나 그 데 와 함께 가 리 - 겨 무
off to see the world. There's such a lot of world to see. We're af -
지 게 풀 따 라 서 - 영 원 한 친구
ter the same rain - bow's end, wait - in' round the bend,
F C Am Dm Gr C
- 내 사 랑 스 러 운 문 리 버 - 와 나
- my Huck - le - ber - ry friend, Moon Riv - er - and me.

수상하였으며, 지금까지 널리 불리고 있다.

이 노래는 1961년 데이비드 멀헤 주제가가 작곡
수상하였으며, 지금까지 널리 불리고 있다.

문 리 버 제 강을 건너 나 그 데 에 게 가 리
Dms Er Am Cr F Bb7
라 - 내 꿈 그 데 만 읊 사 영 하
Am Am/G Am/F# B7 Em7 A7 Dm7 Gr C Am
리 언 제 나 - 그 데 와 - 함께 이 세 상
F C F C Dms Er Am
방 광 하 여 도 나 그 데 와 함께 가 리 - 겨 무
Am/G Am/F# F C F C
off to see the world. There's such a lot of world to see. We're af -
지 게 풀 따 라 서 - 영 원 한 친구
ter the same rain - bow's end, wait - in' round the bend,
F C Am Dm Gr C
- 내 사 랑 스 러 운 문 리 버 - 와 나
- my Huck - le - ber - ry friend, Moon Riv - er - and me.

(a)

(b)

(c)

(d)

+ Sample 2

(a)

(b)

(c)

(d)

Figure 13. Perspective Image is Captured from Camera Galaxy Note 3(a), Staff Correction of Two Top Staves (b), Staff Line Removal of Two Top Staves (c), Output Image on OMR System (d)

3.2 Evaluation

The result of black notes, white notes, whole notes and stem lines have an accuracy of 100% with total of notes. The performance of our recognition system is presented in Table I.

The captured images are corrected staff lines before giving the recognizing music score system. We compare the result of our recognizing symbols method to smart score player (1) and DMSR [1]. If the image is processed, the interface will show and play the music sheet. From the experiments, we release that these failures are caused by noise with high level in the input images.

Table 1. Recognition Results

Symbol recognition	Smart score	DMSR[1]	Our program
Single black note	99.68%	97.12%	98.81%
Multi black note	90.03%	96.8%	96.43%
Single white note	81.85%	95.41%	97.62%
Multi white note	83.76%	95.82%	99.31%
Whole note	78.57%	94.44%	96.65%
Beam	95.5%	95.5%	95.77%
Tail	87.14%	96.2%	99.54%

4. Conclusion

In this paper, we presented a lot of the new approaches to recognize music symbols extracted from the music scores captured by mobile camera with different view point, distort, illumination and noise. To get a desired performance, we do not follow the previous methods in which staff lines are removed, vertical histogram for stem line detection or template matching for musical notes is scanned around stem lines. We remain all staff lines after detecting and restoring its information. With the experiment, it acknowledges music symbols as well as the effective way to create template for matching. We combine staff line removal and without staff line removal so our method shows the highest performance compared with the previous works. Our study is a well-known combination of projects and gets good results. In the future, we would continue to research in this field to recognize entire symbols appearing in the music score and giving these new methods for recognizing other symbols.

Acknowledgments

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (2014-024950. Also this work was partly supported by the Technological innovation R&D program of SMBA [S2173771]. The corresponding author is Guesang Lee.

References

- [1] Vo Q. N., Lee G., Kim S. H. and Yang H. J., "Distorted Music Score Recognition without Staff line Removal," ICPR, August, (2014).
- [2] A. Rebelo, I. Fujinaga, F. Paszkiewicz, A. R. S. Marcal, C. Guedes and J. S. Cardoso, "Optical music recognition: state-of-the-art and open issues," International Journal of Multimedia Information Retrieval, vol. 1, October (2012), pp. 173-190.

- [3] F. Toyama, K. Shoji and J. Miyamichi, "Symbol recognition of printed piano scores with touching symbols," ICPA, (2006), pp. 480-483.
- [4] R. Randriamahefa, J. P. Cocquerez, C. Fluhr, F. Pepin and S. Philipp, "Printed music recognition," Proceedings of the second international conference on document analysis and recognition, (1993), pp. 898-901.
- [5] L. J. Tardón, S. Sammartino, I. Barbancho, V. Gómez and A. Oliver, "Optical Music Recognition for Scores Written in White Mensural Notation," EURASIP Journal on Image and Video Processing, vol. 2009, February (2009).
- [6] B. Su, S. Lu, U. Pal and C. L. Tan, "An Effective Staff Detection and Removal Technique for Musical Documents," Proceedings of the 2012 10th IAPR International Workshop on Document Analysis Systems, (2012), pp. 160-164.
- [7] Jiali C., Huan H. and Yiding W., "An adaptive staff line removal in music score images," ICSP, IEEE, (2010), pp. 964-967.
- [8] M. Droettboom, I. Fujinaga and K. MacMillan, "Optical music interpretation," In IAPR, (2002), pp 378-386.
- [9] P. Soille, "Morphological Image Analysis: Principles and Applications," Springer-Verlag, (1999), pp. 173-174.
- [10] I. Fujinaga, "Staff detection and removal," In George S (ed) Visual perception of music notation: online and offline recognition, Idea Group Inc., Hershey, (2004), pp 1-39.
- [11] G. Taubman, "Musichand: a handwritten music recognition system," In technical report, (2005).
- [12] K. T. Reed and J. R. Parker., "Automatic computer recognition of printed music," In ICPR, vol. 3, (1996), pp 803-807.
- [13] P. Bellini, I. Bruno and P. Nesi, " Optical music sheet segmentation," In Proceedings of the first international conference on web delivering of music, (2001), pp 183-190.
- [14] Sheridan S. and George S., "Defacing music score for improved recognition," In Abraham G, Rubinstein BIP (eds) Proceedings of the Second Australian undergraduate students computing conference, (2004), pp 1-7.

Authors



Van Khien Pham, he received B.S degree from Hanoi University of Sciences and Technology, Vietnam in 2012. Since 2014, he has been taking the M.S. course in Electronics & Computer Engineering at Chonnam National University, Korea. His research interests are mainly in the field of Image Processing, Computer Vision.



Guesang Lee, he received the B.S degree in Electrical Engineering from Seoul National University in 1980. In 1982 He received the M.S degree in Computer Engineering from Seoul National University. In 1991, He received Ph.D. degree in Computer Science from Pennsylvania State University. He is currently a professor of the Department of Electronics and Computer Engineering in Chonnam National University, Korea. Research Interests: Image processing, computer vision and video coding.

