M.S. THESIS

SCENE CLASSIFICATION FOR DEPTH ASSIGNMENT

깊이 정보를 부여하기 위한 이미지 분류

BY

RYU HYUNJOO

FEBRUARY 2016

DEPARTMENT OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

M.S. THESIS

SCENE CLASSIFICATION FOR DEPTH ASSIGNMENT

깊이 정보를 부여하기 위한 이미지 분류

BY

RYU HYUNJOO

FEBRUARY 2016

DEPARTMENT OF ELECTRICAL ENGINEERING AND
COMPUTER SCIENCE
COLLEGE OF ENGINEERING
SEOUL NATIONAL UNIVERSITY

SCENE CLASSIFICATION FOR DEPTH ASSIGNMENT

깊이 정보를 부여하기 위한 이미지 분류

지도교수 김 태 정

이 논문을 공학석사 학위논문으로 제출함

2015년 11월

서울대학교 대학원

전기컴퓨터 공학부

류 현 주

류현주의 공학석사 학위 논문을 인준함

2015년 11월

위 원 장: ______________________________

부위원장: ______________________________

위 원: ______________________________

Abstract

Due to development of 3D display technology, industries related 3D have been grown. For this reason, the demand of 3D contents increases, but there is a short supply of 3D contents. Consequently, research on 2D-to-3D conversion is underway. In 2D-to-3D conversion, the depth information of scene is obtained through an analysis of several depth cues on input sequence and the depth map corresponding to a scene can be generated by combining several depth cues and assigning an appropriate depth level. Scene classification for depth assignment is needed in this process. This paper classifies a scene into landscape, linear perspective, and normal type automatically. The proposed method analyzes landscape type and found there is a relation between image pattern and distribution of color and edge, and suggest the criteria for classification. Moreover, the other criteria for linear perspective classification based on vanishing point detection is proposed. To verify performance, the proposed features are fed into a linear SVM classifier, and 651 images are used. Experiment results show that the algorithm has advantages in performance by about 13%.

keywords: features for scene classification, scene depth assignment, 2D-to-3D conversion, vanishing point detection, distribution of color and edge

student number: 2014-21715

Contents

List of Figures

List of Tables

Chapter 1

Introduction

As industries related to 3D grow, 3D technologies have been developed and the need for 3D content has increased. To generate 3D contents, 2D-to-3D conversion and camera system such as stereo or multi-camera and single camera with depth camera system are required. Using camera system, except 2D-to-3D conversion, is expensive and the supply of 3D content doesn't meet the need for them. With respect to such fact, 2D-to-3D conversion plays an important role [1-3].

Figure 1.1 represents general 2D-to-3D conversion process. In general 2D-to-3D conversion, input sequence is analyzed and a relative depth map is generated according to the scene classification result and depth cues. Next, they control depth level and generate stereoscopic images [5]. Many studies have focused on extracting depth cues and merging them for depth map generation but as shown as in Figure 1.1, scene classification result is used in the overall steps in 2D-to-3D conversion. Consequently, scene classification for depth assigning is helpful in this process.

In this paper, the features extraction method for scene classification is suggested to assign an appropriate depth level. A scene is classified into three type, such as landscape, linear perspective, and normal type [6] and features are extracted according to the type of scene. The extracted features corresponding to scene type outperform compared to the existing method. In the experiment, a linear Support vector machine(SVM) is employed.

The paper is organized as follows. In chapter 2, the concept of scene classification is represented and explained why we need to classify a scene for depth assignment. In chapter 3, the feature extraction method is proposed. That is, features for landscape type are suggested according to scene properties and features of linear perspective type are proposed based on vanishing point detection. Experimental environment and results are explained in chapter 4, and finally conclusion is presented in chapter 5.

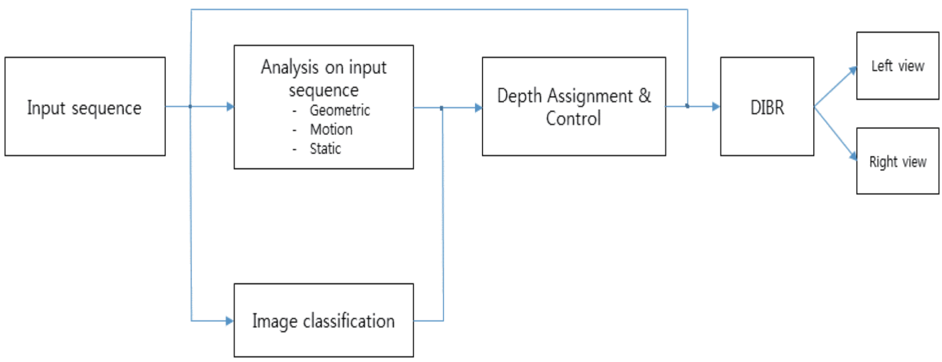

Figure 1.1: 2D-to-3D conversion progress

Chapter 2

Scene Classification for depth assignment

In 2D-to-3D conversion, scene classification has been adopted to get a depth map and there are several methods. One of the famous methods is based on scene geometry [3]. In this paper, a scene is classified into 15 classes, which utilizes color by atmospheric dispersion, geometric context features, texture related features and perspective line features. Next, Lee applies scene classification which is object view class and non-object view class using spectral signatures to generate depth map [4]. Besides, images can be categorized into three classes, such as personal, outdoor, and close-up [7]. In the paper, face detection algorithm, whch utilizes a two-scale wavelet transform for close-up class, and HSI color space for outdoors.

In general, most of studies apply indoor-outdoor classification for depth assignment. S.Battriato employed indoor-outdoor classification results for depth assignment [8]. Besides, Szummer and Picard used color and texture information of each same-sized block and W.Kim adapted edge and color orientation histogram features for indoor-outdoor classification [9-10]. For example, if scene is classified as indoor, small depth level is assigned and vice versa.

This approach is reasonable but there is a limitation in this idea that it could have a small depth level despite outdoor images and the opposite case occurs, too. Figure 2.1 represents examples of limitation on this approach.

Figure 2.1: Limitation case (top) Limitation of outdoor case (bottom) Limitation of indoor case

With the limitation, the other classification method is introduced: a scene is classified into three types, such as landscape type, linear perspective type, and normal type (Figure 2.2) [6]. To generate a depth map, geometric cues like linear perspective and static cues are widely used. When classifying scenes for depth assignment, there is a need to consider those features. That is why a classification into three types is selected. Consequently, a structural depth map plays a role in generating a depth map and Figure 2.3 demonstrates that the method is meaningful with respect to depth assignment.

For example, depth fusion is conducted by combining the structural depth map and the object map.(depth map = α*structural depth map + (1-α)*object map)

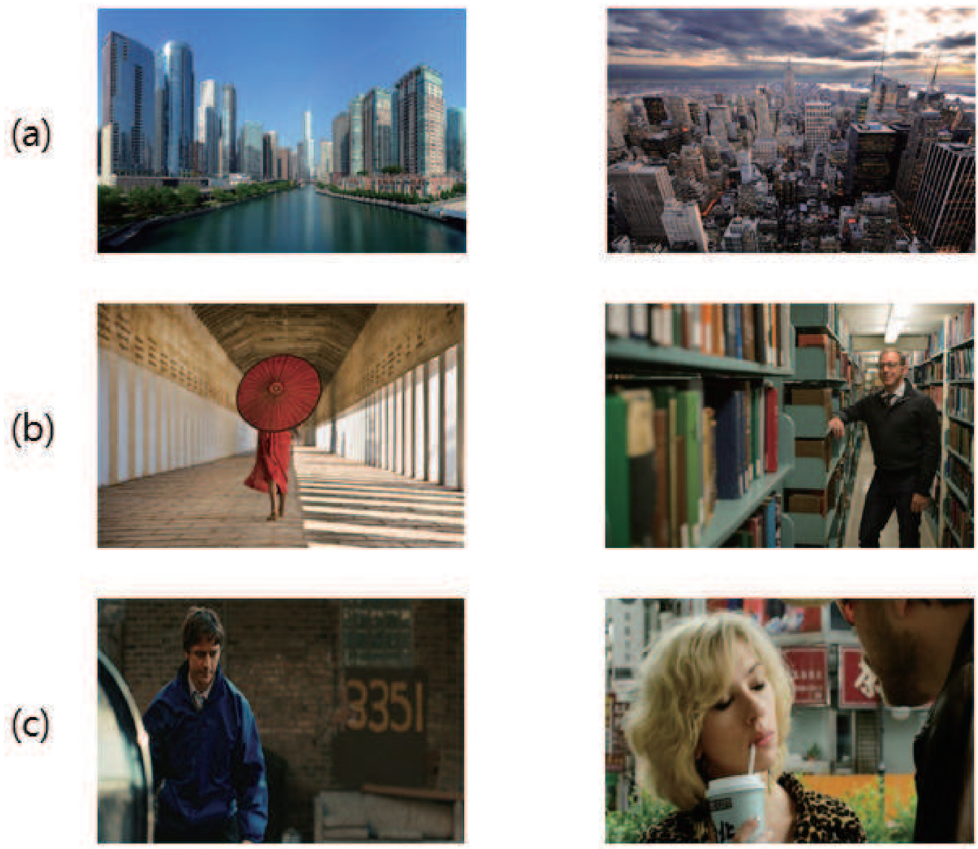

Figure 2.2: Examples of three types (a) Landscape type (b) Linear perspective type (c) Normal type

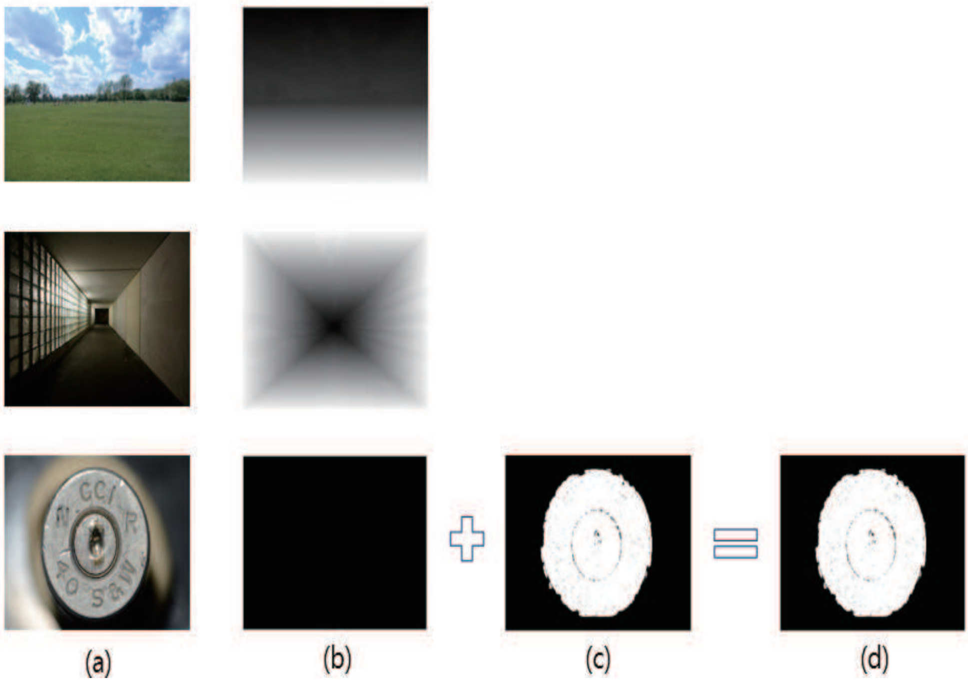

Figure 2.3: Examples of depth map (a) Original image, such as landscape, linear perspective, and normal type above in order (b) Structural depth map (c) Object depth map (d) Final depth map

Landscape type is outdoor scenery image including the sky or the ground region over certain parts. Linear perspective type scene contains the main straight lines which converge to an eventual vanishing point. Other images are the normal type.

First, the proposed algorithm decides the input image whether belongs to landscape type or not. If the image does not belong to landscape type, whether it belongs to linear perspective type is detected. If the image does not belong to the two type, such as landscape and linear perspective type, it is classified as normal type (Figure 2.4).

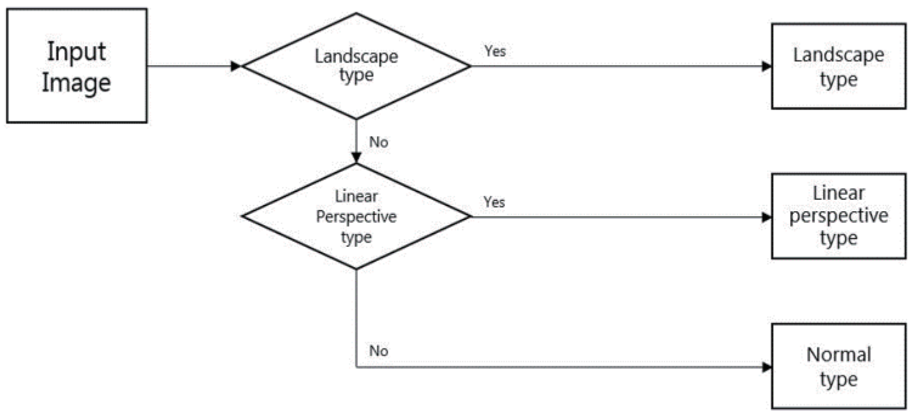

Figure 2.4: Flow chart of scene classification

Chapter 3

Feature Extraction

To classify scene into three types for depth assignment, the criteria for the judgement on each type is needed. We observe properties of each scene type and suggest features corresponding to it : landscape and linear perspective features are proposed.

3.1 Features for landscape classification

3.1.1 Image partition

We observing properties of scene, most landscape images tend to contain many pixels of the sky and the ground regions. For landscape type, the sky locates at the upper part of the image while the ground locates at the lower part. Therefore, the more likely area to get clues for landscape classification will be the upper regions rather than the lower regions. With respect to such fact, a given image is divided into four blocks along the horizontal direction(Figure 3.1). The proposed features are extracted from four divided blocks.

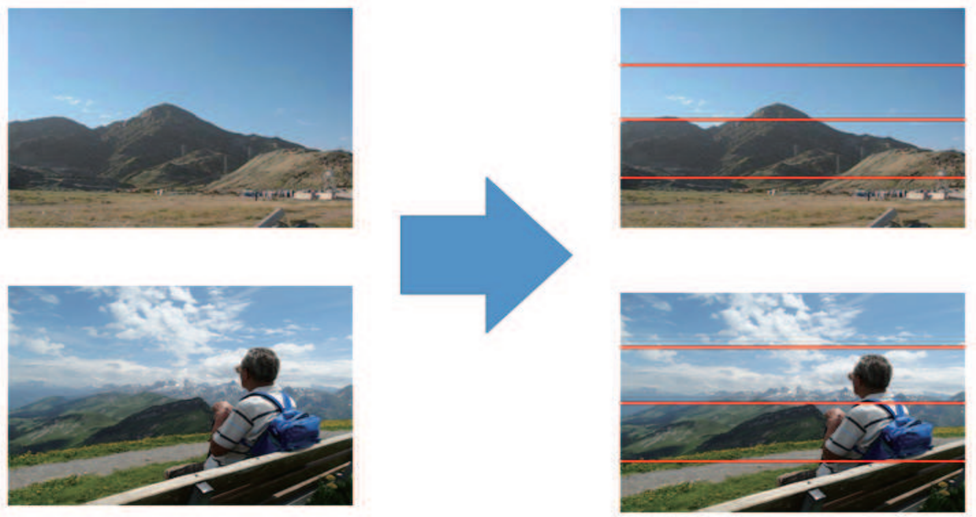

Figure 3.1: Image partition for feature extraction

3.1.2 Color-related features

The relationship between the image and color information is investigated for providing clues of image type. That is, the pixels belonging to the sky or the ground could have certain values within a specific range. For example,

Sky region : H(x,y) $\in$ [100 180] AND I(x,y) $\in$ [100 255].

Ground region : H(x,y) $\in$ [50 100] AND S(x,y) $\in$ [80 255].

H(x,y), S(x,y) and I(x,y) denote the hue, saturation and intensity of each pixel(x,y).

Motivated by this paper, we utilize color information as features for the scene classification, which is mean of the hue of HSI color space, that of intensity, and standard deviation of the hue. Since hue and saturation value infer to the different color tone and reflect the difference between the sky and the ground, we adapt atmosphere scattering features(AS) [3] which is mean and standard deviation of the saturation of HSI model. Also, color orientation 8-bin histogram(COH) [10] is employed.

Although color related features could be one of the significant information for scene classification, it would be inaccurate to use original color. That's because color is strongly affected by illumination color, that is to say color is changed by illumination. Therefore, illumination color estimation for color distribution correction is considered. Color temperature adjustment is commonly operated since D65 denotes average daylight and has a color temperature of approximately 6500K according to the International Commission on Illumination(CIE) [11]. We adapt conditional color temperature adjustment(CCTA) [11] to all color features as mentioned above. Figure 3.2 describes the procedure of illumination color estimation for color distribution correction and Figure 3.3 shows the result of color temperature adjustment.

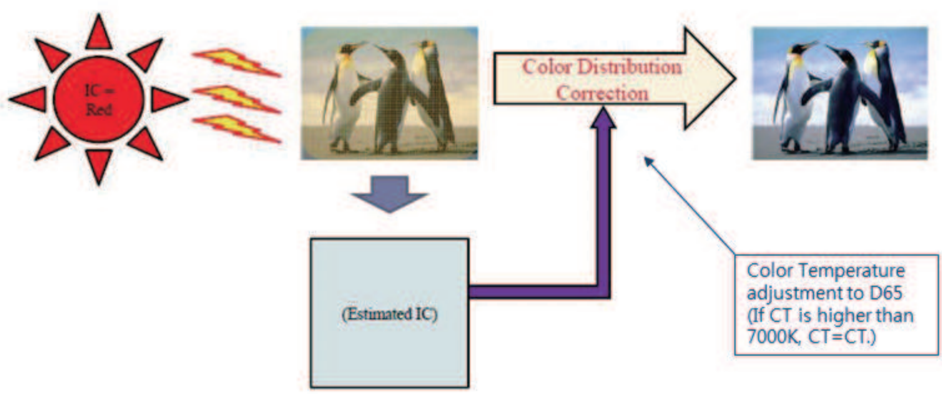

Figure 3.2: Algorithm for conditional color temperature adjustment(CCTA)

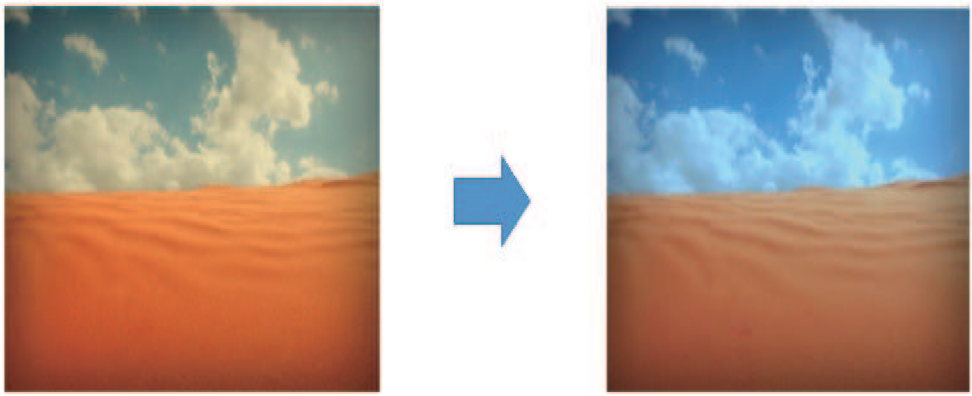

Figure 3.3: The result of CCTA

3.1.3 Edge-related features

In addition to color features, edge information, which presents image structure and properties well, is considered so we focus on edge-related features. There exists a direct relation between landscape image statistics and edge complexity. The ground region of the image tends to be more complex in terms of edge compared to that of the sky region. In [12], the number of edge pixels presents the density of surface and the magnitude of gradient is used to measure the strength of texture. Inspired by this, we utilize gradient sum and edge pixels ratio, which is computed for the sharp edge greater than the threshold. Figure 3.4 depicts the relation between image statistics and the strength of edge.

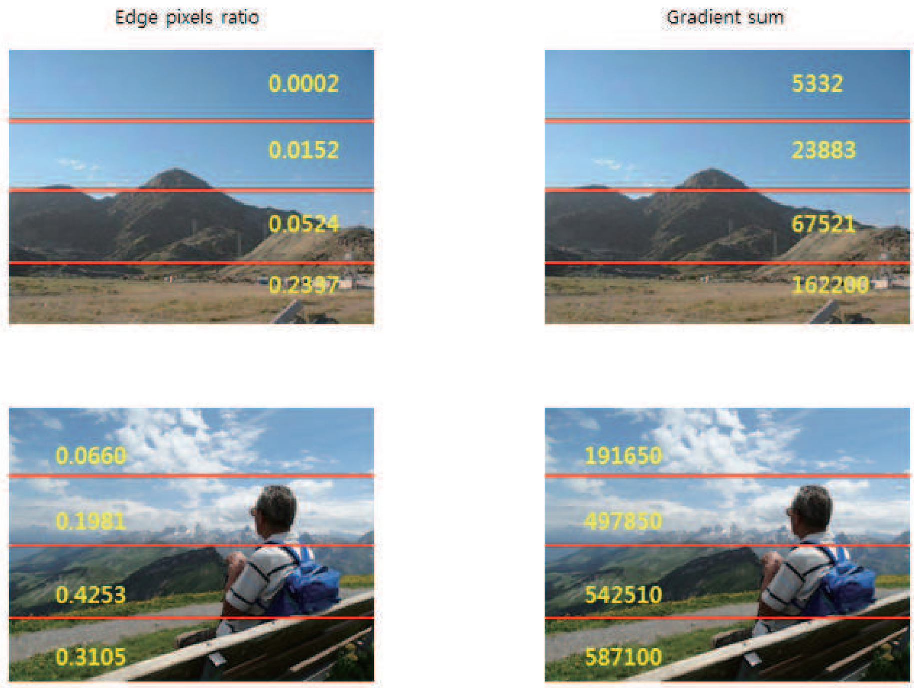

Figure 3.4: Relation between image statistics and strength of edge

Moreover, we observe that edge is more complicated along the horizon compared to the rest so the gradient sum along the horizon per total gradient sum of scene has high value. Because of this, we detect the horizon as follows : edge detection using Sobel and then obtaining the horizon from the highest value along the horizontal direction.

In general, the bottom of images, or the ground region, is relatively more focused and has more high frequency components than the sky region that is the top of images. In [18], there is the relationship between the image and spectral signatures so the textural patterns are represented well by the amplitude spectrum. Thus, the ground can have more pixels more than the average frequency compared to the sky region and it is computed by the amplitude spectrum [19].

In conclusion, we utilize 13 kinds of color and edge features such as mean and standard deviation of hue of HSI color model, 8-bin COH [10], the number of edge pixels motivated from [12], total gradient energy which presents sum of gradient motivated from [12], gradient energy along horizon, AS [3], intensity mean, the number of pixels more than average frequency [19] and COH+AS+mean and standard deviation of hue after CCTA [11].

3.2 Features for linear perspective classification

3.2.1 Criterion for classification of linear perspective type scene

In this paper, we propose the criterion of classification of linear perspective type that is the existence of a vanishing point.

Geometry information in the scene is determined by lines in the scene and the lines extended infinitely cross one point, so called vanishing point. That is, since the lines converge at a vanishing point, linear perspective type is classified by the existence of a vanishing point. Figure 3.5 represents the idea.

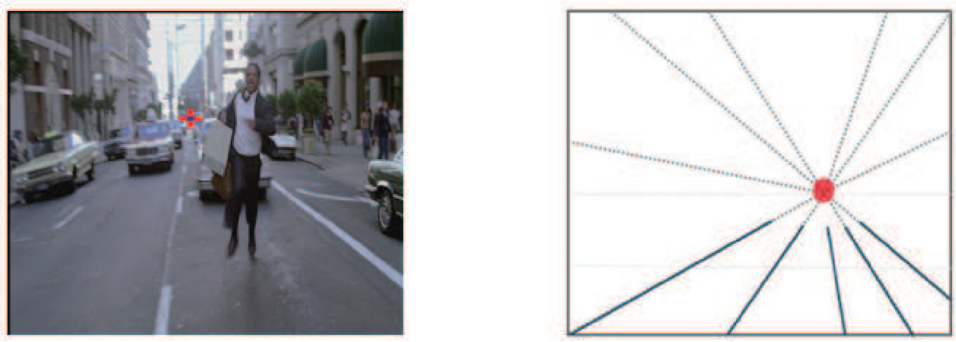

Figure 3.5: Relationship between vanishing point and linear perspective

3.2.2 Vanishing point detection

Vanishing point is defined that straight lines in the image converge on and vanishing lines which cross vanishing point. However, all straight lines in the image don't only pass through vanishing point but also all vanishing lines don't cross at vanishing point. Consequently, detecting a vanishing point accurately plays a important role.

There are several methods for detecting a vanishing point. Rother proposed vanishing detection algorithm which employs camera parameters and find mutual orthogonal direction [13]. It is typical method that Hough Transform is used to detect straight lines in the image and detect intersection point using detected straight lines [14-15]. We focus on this approach and Figure 3.6 describes general vanishing point detection based on Hough Transform approach.

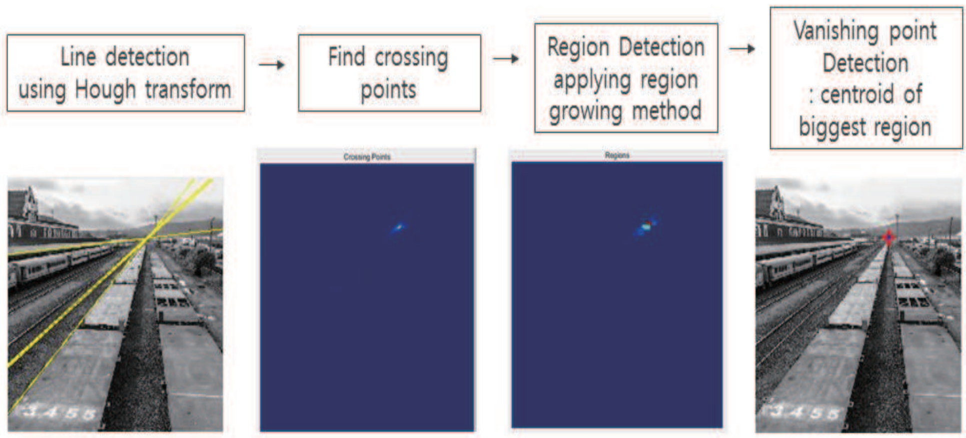

Figure 3.6: General vanishing point detection

Motivated this, we adapt Kim' vanishing point detection algorithm [15] for scene classification in terms of depth assignment. In his paper, a new method of assigning weights to lines is suggested. That is, new lines are constructed by pattern matching on lines detected, and weights to lines are assigned by the ratio of the length of the line segment to the length of the gap. Furthermore, using the region growing method [16], candidate regions likely having vanishing points are determined, and the vanishing point is determined from the centroid of the maximum weight sum of candidate regions (Figure 3.7).

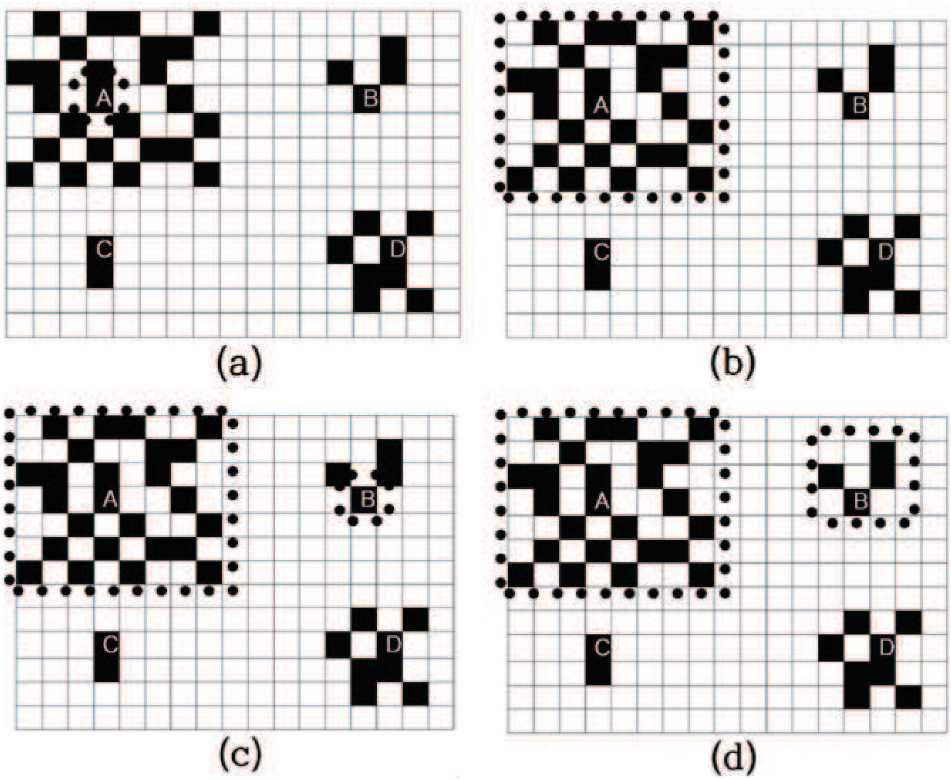

Figure 3.7: Vanishing point detection algorithm of Han (a) The seed as the point having largest weight (b) The selection of the candidate region by region growing method of [16] (c) The selection of next seed (d) The selection of another candidate region

3.2.3 Features based on vanishing point detection

In this paper, we define the criteria for determining linear perspective type is the existence of a vanishing point and the properties for classifying into linear perspective type. It is noted that we apply method of [15] for vanishing point detection.

We can always detect a vanishing point since it assumes a vanishing point always exists in the image. The following characteristics were obtained for the criterion of linear perspective. First, vanishing point exists in the region with the largest weight sum. It could be inferred from the definition of vanishing point that is determined as the centroid of the region with the maximum weight sum (Figure 3.8).

(a) (b)

Figure 3.8: Result images of vanishing point and region map on weight sum (a) Original image with vanishing point (b) Region weight sum map (Blue to Red represents low to high value of weight sum.)

Second, the normalized value of the region with the largest weight sum is the higher proportion on the total weight sum, as shown in Figure 3.8. The normalized value(NV) is computed as follows :

$$NV = \frac{The\ maximum\ weight\ sum}{Total\ weight\ sum} \tag{3.1}$$

Third, the third weight sum besides the largest and second one is considered. In general, there tends to be a great difference between the region with largest weight sum and the region with second weight sum. However, in case scene contains many lines, there is sometimes a small difference despite the existence of a vanishing point. Therefore, the region with third weight sum in addition to the first and second one is considered. Figure 3.9 shows it well.

Figure 3.9: Result images of vanishing point and region map on weight sum (a) Original image with vanishing point (b) Region weight sum map (Blue to Red represents low to high value of weight sum.)

Finally, if vanishing point exists in the scene, the normalized score of the selected region has a high value. Generally, when vanishing point exists in the scene, score of the selected region tends to be high value but score of scene with vanishing point having low value sometimes occurs. In this case, we compensate score by removing noise value with very low, which means we calculate normalized score larger than threshold(Figure 3.10). Normalized score is as follows referencing Figure 3.10 (d).

$$normalized\ score = \frac{average\ of\ a}{average\ of\ T'} \tag{3.2}$$

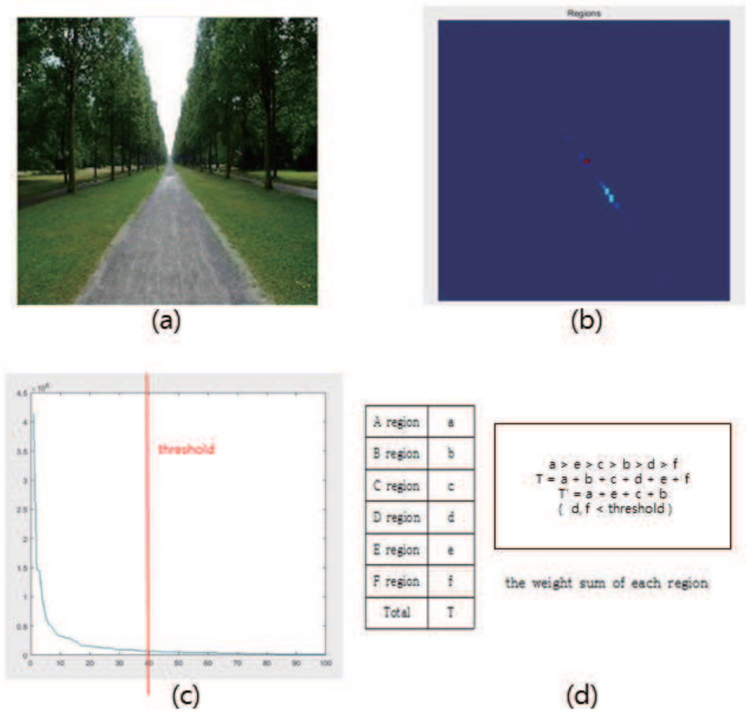

Figure 3.10: Result images of vanishing point and region map on weight sum (a) Original image with vanishing point (b) Region weight sum map (Blue to Red represents low to high value of weight sum.) (c) Plot of region weigh sum map (d) Examples representing the tables and value on the weight sum of each region(A,B,C,D,E,F)

In conclusion, we employ five kinds of features such as the maximum weight sum, total weight sum, the maximum weight sum / total weight sum, (the 1st+ the 2nd + the 3rd weight sum) / total weight sum, and score of the selected region.

Chapter 4

Experiment results

To evaluate the performance of the proposed method, we have conducted experiments on various images collected from several image sequences such as HD movies and photographs. We sorted images into three groups, a landscape scene type group, a linear perspective scene type group, and a normal scene type group. There is a total of 651 images, which consists of 220 landscape group images, 210 linear perspective group images, and 221 normal group images. In the experiment, it is regarded as successful when the result of the classification using the proposed method corresponds to the result of classification sorted by the researchers. Images are resized 480 X 480 and classified by a linear SVM.

4.1 Performance of classification on each step

For landscape type classification, 220 landscape group images, 431 non-landscape group images and total 651 images are tested. In short, 110 landscape images and 216 non-landscape images are trained by linear SVM and the rest of images are classified. We utilize 13 kinds of color and edge features. That is, we extract 116 features such as mean and standard deviation of hue of HSI color model, 8-bin COH [10], the number of edge pixels motivated from [12], total gradient energy which presents sum of gradient motivated from [12], AS [3], intensity mean, the number of pixels more than average frequency and COH+AS+mean and standard deviation of hue after CCTA [11] from four divided blocks. In addition to these features, 3 features related on gradient energy along horizon are extracted from whole image and total dimension is 119.

In terms of performance, Kim's ECOH features [10] are extracted from five divided blocks. Although ECOH is used for indoor-outdoor classification, the reason we chose ECOH is that the landscape / non-landscape classification and indoor-outdoor classification have something in common with the classification criteria for depth assignment. Table 4.1 and Table 4.2 show results of classification.

Table 4.1: Landscape classification results of proposed method

Total accuracy : 90.71% (dimension : 119)

	Landscape	Non-landscape
Landscape	80.91%	19.09%
Non-landscape	4.19%	95.81%

Table 4.2: Landscape classification results of existing method(ECOH) from [10]

Total accuracy : 75.69% (dimension : 80)

	Landscape	Non-landscape
Landscape	71.82%	28.18%
Non-landscape	22.33%	77.67%

Apply the proposed method, the total accuracy is 90.71% while that of the existing method is 75.69% then the classification rate increases from 75.69% to 90.71%.

As for the result, there is a limitation of ECOH, in which the concept of classification differs in the way the scene with less sky region among other outdoor images are regarded as non-landscape type by the proposed method. Figure 4.1 depicts the difference in concept.

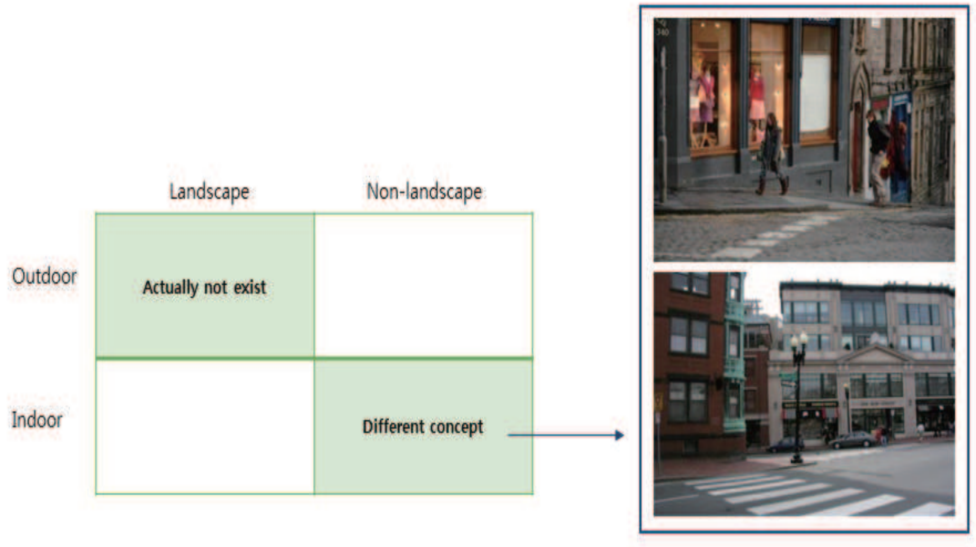

Figure 4.1: The difference in concept between proposed method and existing method

In addition to the dissimilarity in the concept, ECOH features are focused on orientation of edge but the proposed method is considered the strength of edge according to each divided blocks such as edge pixels ratio, gradient sum, gradient sum along the horizon per total gradient sum, and the number of pixels more than average frequency. It is meaningful because the ground region of the scene tends to be more complicated in terms of edge compared to that of the sky region. Moreover, the proposed method utilizes color-related features along with COH such as mean of hue, mean of intensity, and standard deviation of hue. AS from [3] as well as hue and intensity related features reflects the characteristics of scene with respect to color tone of ground and sky region. Furthermore, there is effective by adjusting color temperature when color is changed by illumination. As a result, the proposed method for landscape type classification outperforms the existing method, ECOH.

For the validation of linear perspective type classification, 210 linear perspective group images, 441 non-linear perspective group images, and the total 651 images are used. So, we use 105 linear perspective images and 221 non-linear perspective images to train with a linear SVM and the rest to classify.

Five features based on vanishing point detection such as the maximum weight sum, total weight sum, the proportion of the maximum weight sum to total weight sum, the rate of the maximum, the 2nd, and the 3rd weight sum to total weight sum, and the normalized score of selected region are extracted for classification. As compared with the proposed method, Hao's method [6] is introduced where it is based on general line detection using Hough transform. It is composed of 5 features that are voting value threshold, the number of lines along voting value threshold, total voting value for line detection, maximum voting value for line detection, and the ratio of maximum voting value to total voting value. Table 4.3 and 4.4 present results of classification.

Due to sophisticated algorithm for detecting a vanishing point, the proposed method shows the better classification results compared to the existing method [6], which the total accuracy rises from 70.46% to 82.46% despite same dimension.

Table 4.3: Linear perspective classification results of proposed method

Total accuracy : 82.46% (dimension : 5)

	Linear perspective	Non-linear perspective
Linear perspective	72.38%	27.62%
Non-linear perspective	12.73%	87.27%

Table 4.4: Linear perspective classification results of existing method from [6]

Total accuracy : 70.46% (dimension : 5)

	Linear perspective	Non-linear perspective
Linear perspective	89.52%	10.48%
Non-linear perspective	38.64%	61.36%

4.2 Performance of scene classification result for depth assignment

In the experiment, 110 landscape group images, 105 linear perspective group images and 111 normal group images are used to train with a linear SVM and the rest of the 651 total images consisting of 110 landscape, 105 linear perspective, 110 normal images are tested. We conducted the experiment through 2-step classification. First, the scene is judged whether it is a landscape type or not, and then the rest is determined by linear perspective and normal type.

In terms of performance, the proposed method features mentioned in the previous section are extracted and we adapt ECOH from [10] for the 1st step classification, landscape type classification, and Hao's method from [6] for the 2nd step classification, linear perspective type classification, as also remarked in the previous section.

Table 4.5 and Table 4.6 show results of whole classification. Applying the proposed method, the accuracy of the classification is 80.91%, 73.33%, and 95.45% on landscape, linear perspective, normal type, respectively. Meanwhile, the accuracy of the classification adapting the existing method is 71.82% on landscape, 70.48% on linear perspective, and 65.45% on normal type. Finally, the total classification rate increases from 69.23% to 83.38% so the proposed method is fairly reasonable.

Table 4.5: Scene classification results of proposed method

Total accuracy : 83.38% (dimension : 119 / 5)

	Landscape	Linear perspective	Normal
Landscape	80.91%	7.27%	11.82%
Linear perspective	7.62%	73.33%	19.05%
Normal	0.91%	3.64%	95.45%

Table 4.6: Scene classification results of existing method combined with [10] and [6]

Total accuracy : 69.23% (dimension : 80 / 5)

	Landscape	Linear perspective	Normal
Landscape	71.82%	6.36%	21.82%
Linear perspective	20.95%	70.48%	8.57%
Normal	23.64%	10.91%	65.45%

Figure 4.2 demonstrates scene classification into 3 type is valid with respect to depth assignment. The model for generation of depth map referenced from [17] is as follows.

$$D_{fusion} = \alpha D_{structural} + (1 - \alpha) D_{object} \tag{4.1}$$

D_{fusion}, $D_{structural}$ and D_{object} denote final depth map, structural depth map and object depth map, respectively.

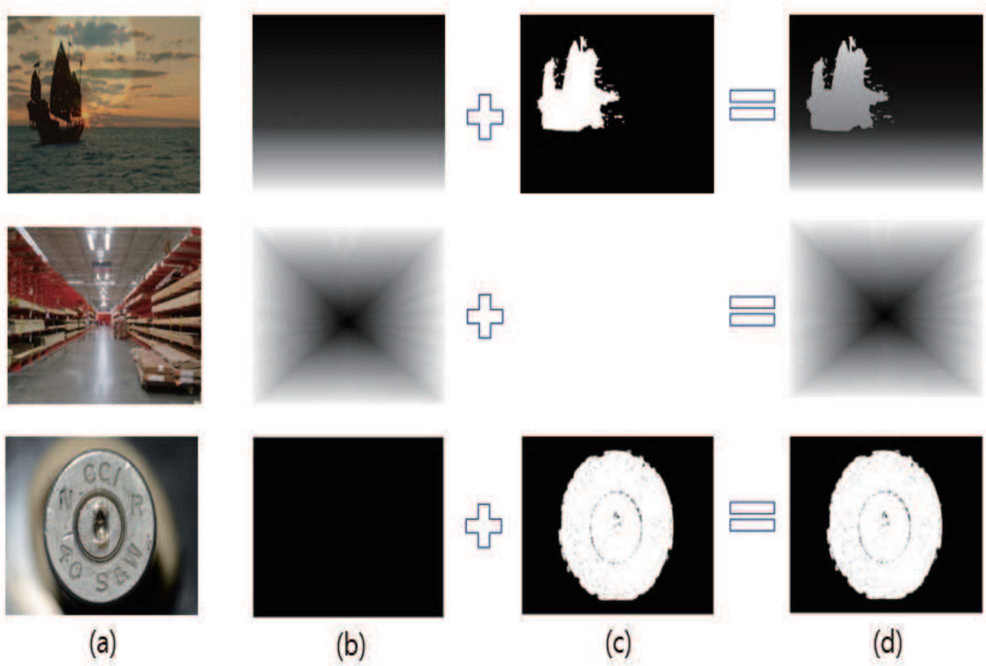

Figure 4.2: Generation of depth map according to scene classification (a) Original image (b) Structural depth map (c) Object depth map (d) Final depth map

Chapter 5

Conclusion

In this paper, we propose an automatic scene classification algorithm for depth assignment. Scene is classified into three type, landscape, linear perspective, and normal type, and features are extracted in accordance with scene characteristics through analysis. Features for three kinds of scene classification are presented. For landscape classification, we found that a direct relation between color and edge pattern and landscape type scene and introduced features related it. Also, for linear perspective classification, new features based on vanishing point detection are suggested.

To validate the performance of the proposed method, various images have been tested and proposed features are fed into linear SVM classifier. As a result, the proposed method outperforms compared with existing method and is also fairly meaningful with respect to depth assignment.

The proposed method can be used for depth control for generating depth map and more realistic 3D contents. Further, out method is very useful for real-time application.

Bibliography

[1] Feng, Yue, Jinchang Ren, and Jianmin Jiang, ”Object-based 2D-to-3D video conversion for effective stereoscopic content generation in 3D-TV applications.” *IEEE Transactions on Broadcasting*, vol. 57, no.2, pp. 500-509, June 2011.

[2] Zhang, Liang, Carlos Vazquez, and Sebastian Knorr, ”3D-TV content creation: automatic 2D-to-3D video conversion.” *IEEE Transactions on Broadcasting*, vol. 57, no. 2, pp. 372-383, June 2011.

[3] V. Nedovic, A. w. Smeulders, A. Redert, J. M. Geusebroek, ”Stages as models of scene geometry.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 32, no.9, pp. 1673-1687, September 2010.

[4] J. Lee, S. Yoo, C. Kim, B. Vasudev, ”Estimating scene-oriented pseudo depth with pictorial depth cues.” *IEEE Transactions on Broadcasting*, vol. 59, no. 2, pp. 238-250, June 2013.

[5] Caviedes, Jorge, and Javier Villegas, ”Real time 2D to 3D conversion: Technical and visual quality requirements.” in *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 897-898, January 2011.

[6] Hao Dong, Shouyi Yin, Guangli Jiang, Leibo Liu and Shaojun Wei, ”An automatic depth map generation method by image classification.” in *IEEE International Conference on Consumer Electronics (ICCE)*, pp. 168-169, January 2015.

[7] Huang, Yea-Shuan, Fang-Hsuan Cheng, and Yun-Hui Liang, ”Creating depth map from 2D scene classification.” in *IEEE 3rd International Conference on Innovative Computing Information and Control (ICICIC)*, pp. 69-69, June 2008.

[8] S. Battiato, S. Curti, M. La Cascia, M. Tortora, and E. Scordato, ”Depth map generation by image classification.” in *International Society for Optical Engineering on Three-Dimensional Image Capture and Applications VI*, vol. 5302, pp. 95-104, January 2004.

[9] Szummer, Martin, and Rosalind W. Picard, ”Indoor-outdoor image classification.” in *IEEE International Workshop on Content-Based Access of Image and Video Database, Proceedings*, pp. 42-51, January 1998.

[10] W. Kim, J. Park, and C. Kim, ”A novel method for efficient indoor-outdoor image classification.” in *Journal of Signal Processing Systems*, vol. 61, no. 3, pp. 251-258, December 2010.

[11] 갈준호, 김보은, 송원석, 김태정, ”이미지 분류의 기능 향상을 위한 색온도 조정 기반의 보정 기법.” *대한전자공학회 2013년 하계학술대회*, 제36권, 제1호, 2013년 7 월. pp. 1892-1895.

[12] 갈준호, 최두섭, 송원석, 김태정, ”2차원 영상의 3차원 변환에서 깊이 정보 부여를 위한 이미지 분류.” *대한전자공학회 2012년 추계학술대회*, 2012년 11월. pp. 521-524.

[13] C. Rother, ”A new approach to vanishing point detection in architectural environments.” *Image and Vision Computing*, vol. 20, no. 9, pp. 647-655, 2002.

[14] V. Cantoni, L. Lombardi, M. Porta, and N. Sicard, ”Vanishing point detection: representation analysis and new approaches.” in *11th IEEE International Conference on Image Analysis and Processing*, pp. 90-94, 2001.

[15] 김항태, 송원석, 최혁, 김태정, ”인공 구조물 내 직선을 찾기 위한 경험적 가중치를 이용한 소실점 검출 기법.”, *정보과학회논문지*, 제42권, 제5호, 642-651쪽, 2015년 5월.

[16] J. Han, ”Relative Depth-Map Generation from Monocular Image using a Valid Vanishing Point.”, in *Proc. of Korea Computer Congress*, vol. 33, no. 2, 2006.

[17] Jaeseung Ko, Manbae Kim, and Changick Kim, ”2D-to-3D stereoscopic conversion: depth-map estimation in a 2D single-view image.” in *International Society for Optics and Photonics on Optical Engineering+ Applications*, pp. 66962A-66962A-9, 2007.

[18] A. Torralba, and A. Oliva, ”Depth estimation from image structure.” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no.9, pp. 1226-1238, November 2002.

[19] Xiaodi Hou, and Zhang Liqing, ”Saliency detection: A spectral residual approach.” in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR’07)*, pp. 1-8, June 2007.

국문 초록

3D 관련 디스플레이 기술이 발전하면서 관련 산업에 대한 관심이 높아지고 시장 역시 규모가 커지고 있다. 이에 3D 컨텐츠의 수요는 증가하였으나, 공급이 그에 미치지 못하면서 2D 영상을 3D로 변환하는 방법에 대한 연구가 활발히 진행되고 있다. 2D 영상을 3D로 변환할 때에는 여러가지 단서들을 이용하여 깊이 정보를 얻고 해당 장면에 맞게 적절한 수준의 깊이를 부여하여 깊이 지도를 만들게 된다. 이 과정에서 2D 영상에 대한 분류를 수행하는 것은 효과적으로 깊이를 부여하기 위해서 필요하다. 본 논문에서는 효과적인 깊이 부여를 위해 영상 분류를 할 때 필요한 특징 추출 방법을 제안하였다. 자연 현상 관찰을 통하여, 영상의 분포에 따른 색상과 엣지 분포의 관계를 발견하고 그에 맞는 특징을 추출하였으며, 선 원근법과 깊이 정보와의 관계를 이용하여 소실점 추정을 이용한 선 원근법 관련 특징을 제안 하였다. 제안된 특징들은 효과적인 깊이 부여을 위한 영상 분류의 특징을 잘 표현하였다. 마지막으로 제안된 특징들을 이용하여 서포트 벡터 머신을 통한 영상 분류의 성능을 실험 하였다. 제안된 특징 추출 방법으로 성능을 확인하였다.

주요어: 이미지 분류를 위한 특징, 효과적인 깊이 부여, 2D 영상의 3D 영상 변환, 소실점 추정, 색상과 엣지 분포

학번: 2014-21715