A New Similarity Measurer for Color-Texture and Its Clustering for Apple

Jun-Hyeok Lee, Ji-Hong Park, and Seon-Jong Kim
Dept. of IT Engineering, Pusan National University, Miryang, Kyeongnam, South Korea
Email: {lllalll2, qkrwlghddlek, ksj329}@pusan.ac.kr

***Abstract*—This paper proposed a new similarity index for applying to clustering by color-texture using the IOCLBP. We evaluated the relationships of IOCLBP component pairs and investigated a distribution of color-texture properties for apples. And then, we found some correlations between their properties. As selecting the uncorrelated properties and their centric feature vector of apples, the proposed similarity index was introduced as a distance measurer. In experiments of clustering with apple fruits, the results showed that the proposed index has a good performance and is used in color-texture measurer.**

***Index Terms*—color-texture, similarity measurer, apple proper-ties, clustering**

I. Introduction

Human can acquire many features of an object by the owned eyes, and use these features to get some information about the object. Shape, color, and texture are mainly used in representing the features with a measurer. For example, the human eye can be used effectively in classifying textures, such as rough, soft, or irregular. However, there is ambiguous about their quantitative value that is universally recognized and accepted. The various definitions have been proposed in many studies in fields of computer vision.

It is important to represent a degree of the shape, color, and texture in the given image. LBP (local binary pattern) have been widely studying for texture descriptors due to having a good performance [1], [2]. The initial LBP are used with a gray image. There also have been studying a lot of efforts to integrate color and texture information into a combined descriptor recently [3].

Especially, an IOCLBP (improved opponent color LBP) shows a good performance as a descriptor that combine the color with the texture features. The IOCLBP can make feature vectors to lower dimension by applying an invariant to rotation. It is useful for obtaining the specific patterns depending on local neighboring pixels.

However, IOCLBP tends to be generated the redundant high dimensional features that may not be usefulness for clustering in multiple color spaces. And the specific objects could be inherited from the similar properties in shape, color, or texture. With that case, we should reduce a feature space with some extent and use the inherent sparsity of the features itself for doing this [4]. High dimensional features with redundancy can be found in most with same. Therefore, we should establish the relation with low dimensional to be useful for quantitative measurer of the similarity.

In this paper, we propose a new similarity index using the IOCLBP, and apply it to clustering color-texture of apples. We evaluate IOCLBP component pairs based on color-texture for apples firstly. And we investigate the relationship of the given properties. We also define the centric feature vector about each color and introduce an texture index as a distance from this vector away.

II. Similarity Index for Color-Texture

A. Rotation-Invariant LBP

As mentioned before, LBP have been widely using for texture descriptor. In LBP, each pixel has 0 or 1 by comparing with the center pixel of the 3x3 array as shown in Fig. 1. For the given local pixels of (a), 172 value at the center pixel is finally changed to 162 by the neighborhood pixels in LBP. LBP can be used effectively in representing the local surrounding features.

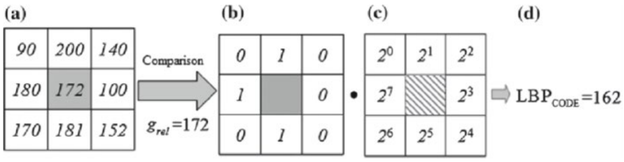


Figure 1. The principles of the basic LBP.

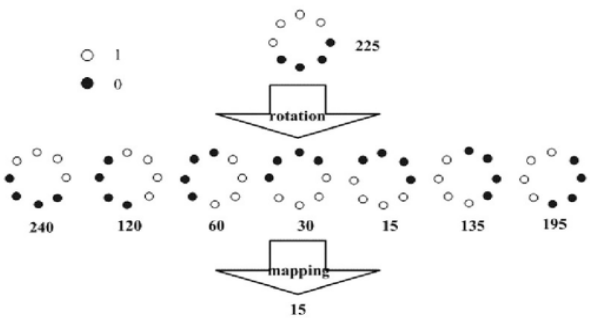


Figure 2. The value of 225 is the same as 15 in case of rotation-invariant LBP.

This basic LBP is a robustness even if brightness is changed. Also with an invariant to rotation of the patterns, we can map out 225 to 15 as shown in Fig. 2. That is, the value of 225 is the same as 15 if it is rotated. It also has

Manuscript received March 29, 2019; revised August 18, 2020.

doi: 10.18178/ijsps.8.3.59-63

the same value of 15 as 240, 120, 60, 30, 135 and 195. The active ranges can reduce to 36 in case with the rotation-invariant LBP. They remain a value with 0, 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 37, 39, 43, 45, 47, 51, 53, 55, 59, 61, 63, 85, 87, 91, 95, 111, 119, 127, and 255. Fig. 3 is shown an example of LBP histogram with rotation invariants in the given image. It depends on the image whether the value exists in histogram.

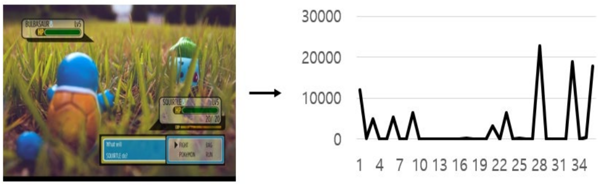


Figure 3. An example of histogram with rotation invariant LBP.

B. IOCLBP

There are many methods to use both color and texture. It can be presented by two directions. One is to separate color from texture, and the other is to consider both together. IOCLBP can use the components as other color space to compare with local value of neighbor, and is belonged in the latter [6].

For (*u, v*) color space, IOCLBP is defined as follows;

$$f_{IOCLBP_{u,v}}(P) = \sum_{i=0}^{n} 2^i \Phi(\overline{p}_u, p_{i,v}) \quad (1)$$

$$\overline{p}_u = \frac{1}{n}\sum_{i=0}^{n} \overline{p}_{i,u}$$

where, $p_{i,v}$ indicates an intensity of the center pixel of *v* color space and $\overline{p}_u$ indicates an average intensity of the i_th neighbor pixel of u color space. The function ϕ() has a 0 or 1 according to LBP. And m means the number of local neighbor, is 8. We use four color components, generally R, G, B and Gray [7]. We know it is totally 16 component pairs to apply for IOCLBP.

For the given image, we can calculate a histogram of IOCLBP image by each component pairs. Therefore, an IOCLBP histogram T consists of as follows;

$$T = [A_1, A_2, A_3, \ldots, A_n] \quad (2)$$

$$A_i = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1q} \\ a_{21} & a_{22} & \ldots & a_{2q} \\ \ldots & \ldots & \ldots & \ldots \\ a_{p1} & a_{p2} & \ldots & a_{pq} \end{bmatrix}$$

where, n means the number of objects, p means the number of components, is 16, and q means the range of LBP, is 36. That is, it can obtain 16 histogram bars with 36. For size, the histograms are normalized as follows;

$$a_{jk} = \frac{n_{jk}}{\sum_l^p n_{ljk}} \quad (3)$$

C. A Centric Vector

Some objects have a similar value of color and/or texture [8]. Sometimes, it is useful if we know an average color or texture of the species or class. It is difficult to get information about all of them. We can obtain them roughly from a part. We call it a centric vector C as follows;

$$C = \begin{bmatrix} c_{11} & c_{12} & \ldots & c_{1q} \\ c_{21} & a_{22} & \ldots & c_{2q} \\ \ldots & \ldots & \ldots & \ldots \\ c_{p1} & c_{p2} & \ldots & c_{pq} \end{bmatrix} \quad (4)$$

$$c_{jk} = \frac{1}{p}\sum_{i=1}^{p} T_{ijk}$$

A centric vector means an average vector of the all sampled color or texture space as equation (4). Most of feature vectors of the objects exist around of this centric vector due to a similar color or texture. This means that the color and the texture do not deviate on large scale from the inherent of the given class.

D. Similarity Index

A centric vector can simply be used in a distance measurer. For some species, there is a pattern that comes out mostly. The probability of this pattern is an important role for determining a distance from the centric vector. We define a new similarity index based on the fact that the distance form centric vector is related to the probability of the primary pattern. An index of similarity *Si* with a given vector *Ai* is defined as follows;

$$S_i = 1 - \frac{A_i \cdot C}{C \cdot C} \quad (5)$$

Dot means the inner products. The inner product indicates a difference between the vectors.

III. Experiment Results

A. Dataset

Fig. 4 showed a part of the apple images used in our dataset. We cut the images as shown in Fig. 5. We analyzed color-texture properties of this regions that has a regular pattern for color-texture. The stalk parts is excluded because it has rather other features than color [9]. We collected the images from 120 apples.

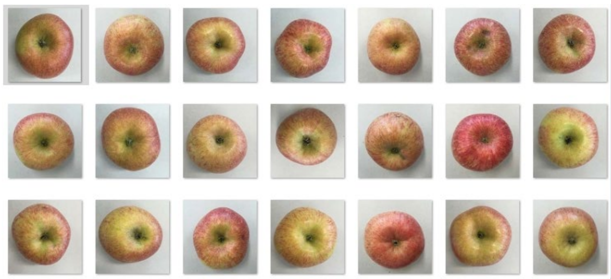


Figure 4. A part of our dataset with apple images.

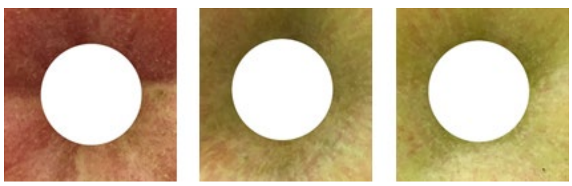


Figure 5. Images for color-texture.

B. Histogram Analysis

It is clear that histograms have a distribution of 36 values with a rotation-invariants. Through the experiments, we found that histograms of IOCLBP in apples consist of just 6 specific patterns of more than 97%. To confirm, Fig. 6 is shown the sum of total IOCLBP patterns. In most case, it showed that just 6 patterns showed up.

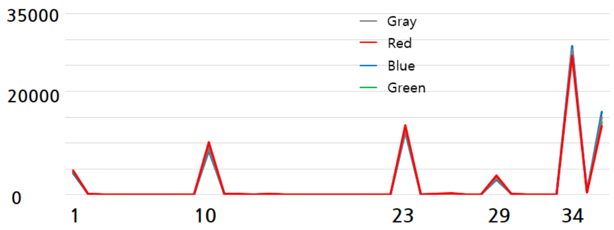


Figure 6. A one of IOCLBP histograms in 4 color spaces.

For example, the 36 LBPs was totally 8,914,581 and among them it was 8,698,670 indicating just 6 patterns in case of Gray image. It accounts for 97.58%. We can check that the rests of this patterns are very irregular and small. It appears in similar to red, blue, and green too.

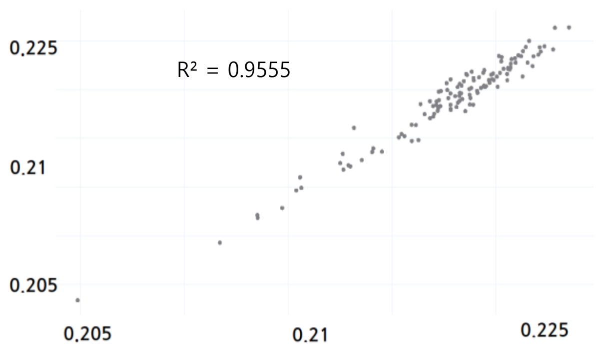
C. Component Pairs Analysis

Before applying our proposed similarity index, we must evaluate the relationships between component pairs to reduce the redundancy. For (*u, v*) space of IOCLBP, we can select the four color. Total components of R, G, B and Gray are shown in Table I. *w* means the Gray space.

TABLE I. IOCLBP COLOR WITH GRAY SPACE ADDED

Color space	R	G	B	w
R	RR	RG	RB	Rw
G	GR	GG	GB	Gw
B	BR	BG	BB	Bw
w	wR	wG	wB	ww

Lower matrix (GR, BG, BR, wR, wG, and wB) are a reverse cases of upper matrix. They are redundant. This is why they have the contrary value of 0 and 1. We have also confirmed that the relationships had nearly linear between the component pairs of RR, GG, BB, and ww each other, as shown in Fig. 7. The linearity of Gw space is small, but we supposed to linear relationship due to its weakness and visualization.



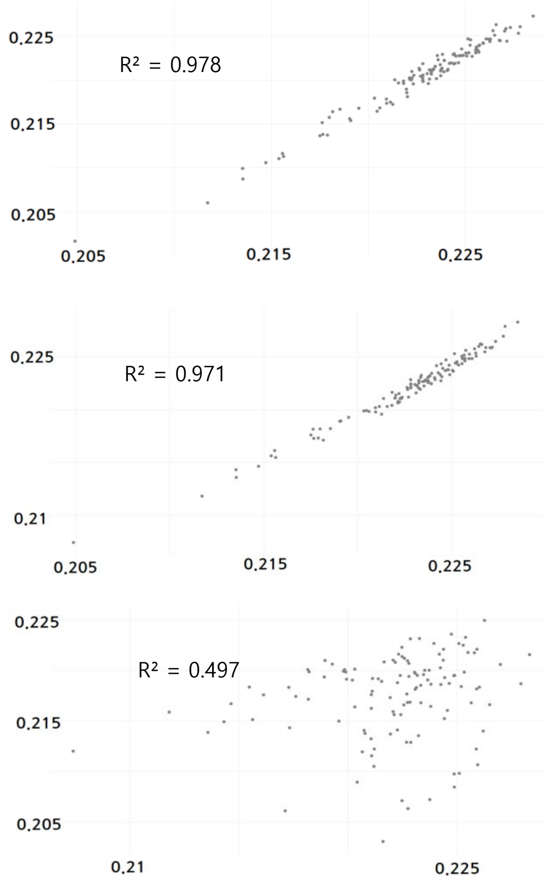


Figure 7. Relationship between the color spaces of (RR, GG), (RR,BB),(RR, ww) and (RR, Gw).

And we found that Bw, RB, and GB have a nearly linear relationships each other as shown in Fig. 8.

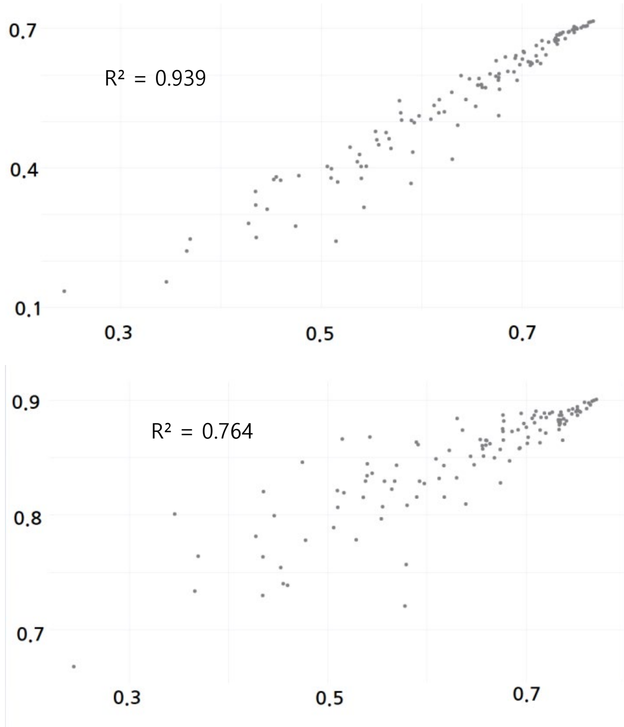


Figure 8. Relationship between the color spaces of (Bw, RB) and (Bw,GB).

We finally confirmed that Rw and RG have a linear relationships as shown in Fig. 9.

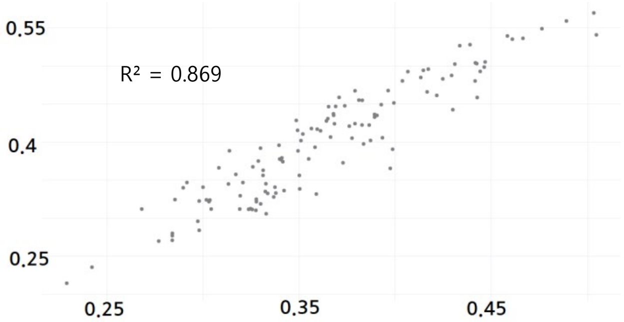


Figure 9. Relationship between the color spaces of (Rw, RG).

As a result, we can just find the component pairs of 3 classes, and then use 3 components pairs independently. We also investigated the correlation between RR, Rw, Bw, and Gw respectively, shown in Table II. RR, Rw, and Bw with low scores were selected in our experiments for more correlation [10].

TABLE II. CORRELATION RR, RW, BW, GW

Correlation	RR	Rw	Gw	Bw
RR		0.015	0.15	-0.29
Rw			-0.4009	0.056
Gw				0.64
Bw				

D. Centric Vectors

Fig. 10 is showed centric vectors of RR, Rw, and Bw respectively. This is an average vector of the normalized histograms for 120 apples. We knew that pattern 1 appeared very often in example of Bw. This is the primary pattern in our apple colors.

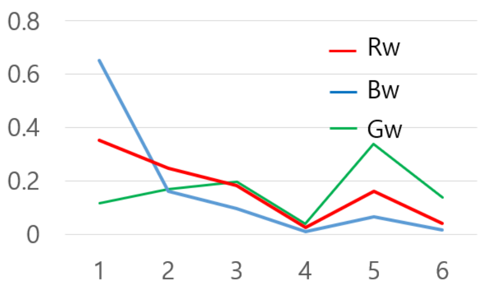


Figure 10. A centric vector of Rw, Bw and Gw components.

Fig. 11 is depicted an example of histogram in RR of two apples. The centric vector showed up between two vectors. Since the proposed similarity index uses as a distance measurer from the centric vector away, the similarity index is large in case that the primary pattern appeared frequently indicted in orange color in Fig. 11. Or the similarity index is small in case that the primary pattern appeared sometimes indicated in blue color.

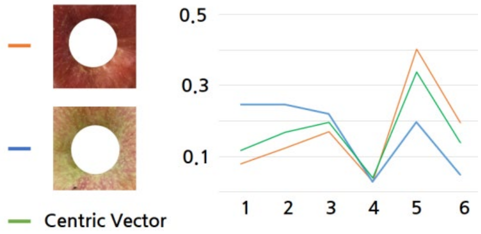


Figure 11. The example of two apple vectors and a centric vector in RR component pairs.

E. Color-Texture Clustering Using K Means

We experimented a clustering of the given apples by the proposed similarity index, that is, by a color-texture similarity measurer. As you know, our spaces have only 3 component pairs, RR, Rw, and Bw. We used k means algorithm to clustering apples with our 3 similarity index in our experiments. Using the proposed similarity index, we can represent in 3D spaces, and 120 apples were divided into k=60 clusters as shown in Fig. 12.

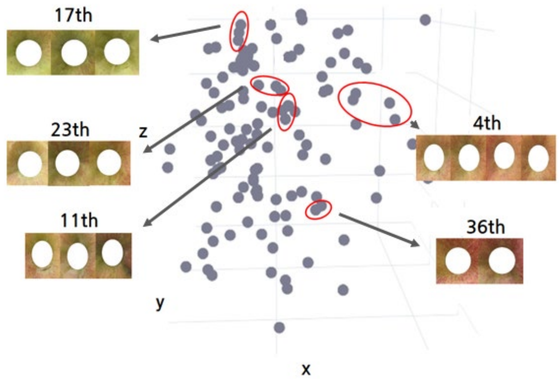


Figure 12. The 3D visualization of 120 Apples using the proposed 3 pairs of the proposed similarity index.

IV. CONCLUSIONS

This paper proposed a new similarity index for applying to cluster color-texture of apples using the IOCLBP. We evaluated IOCLBP component pairs for apples, investigated a distribution of color-texture properties. And we found 3 component pairs. As defining the centric feature vector, the similarity index was proposed as a distance measurer. From experiment results, we can reduce 16 spaces to three by the linearity between the component pairs, RR, Rw, and Bw. With 3 pairs, we can visualize the clustering results in 3D spaces. Clustering results of apples showed the proposed index have a good performance as a color-texture measurer.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Seon-Jong Kim, Jun-Hyeok Lee conducted the research; Ji-Hong Park analyzed the data; Seon-Jong Kim, Jun-Hyeok Lee wrote the paper; all authors had approved the final version.

ACKNOWLEDGMENT

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2018R1D1A1B07045565).

REFERENCES

[1] A. Porebski, V. T. Hoang, N. Vandenbroucke, and D. Hamad, "Multi color space LBPbased feature selection for texture classification," *Journal of Electronic Imaging*, vol. 27, no. 1, p. 011010, 2018.

[2] V. T. Hoang, "Multi color space LBP-based feature selection for texture classification," PhD thesis, Université Lille Nord-De-France, 2018.

[3] S. Banerji1, A. Verma, and C. Liu, "New image descriptors based on color, texture, shape, and wavelets for object and scene image classification," *Neurocomputing*, vol. 117, pp. 173-185, Oct. 2013.

[4] F. Bianconi, R. B. Cerezo, P. Napoletano, and F. D. Maria, "Improved opponent color local binary patterns for color texture classification," *Lecture Notes in Computer Science*, pp. 272-281, March 2017.

[5] J. S. Park, "An effective algorithm for subdimensional clustering of high dimensional data," *Journal of The KIPS Transactions: Part D*, vol. 10D, no. 3, pp. 417-426, Jan. 2003.

[6] F. Bianconi, R. B. Cerezo, and P. Napoletano, "Improved opponent color local binary patterns: An effective local image descriptor for color texture classification," *Journal of Electronic Imaging*, vol. 27, no. 1, Feb. 2018.

[7] S. Banerji, A. Verma, and C. Liu, "Novel color LBP descriptors for scene and image texture classification," in *Proc. International Conference on Image Processing, Computer Vision, and Pattern Recognition*, July 2011, pp. 1-7.

[8] E. G. Ortiz, A. Wright, and M. Shah, "Face recognition in movie trailers via mean sequence spare representation-based classification," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, September 2013, pp. 3531-3538.

[9] S. R. Dubey and A. S. Jalal, "Fusing color and texture cues to categorize the fruit diseases from images," *International Journal of Computer Vision and Image Processing*, vol. 4, no. 2, pp. 52-67, 2014.

[10] M. Kalakech, A. Porebski, N. Vandenbroucke, and D. Hamad, "Unsupervised local binary pattern histogram selection scores for color texture classification," *J. Imaging*, vol. 4, no. 10, p. 112, Sept. 2018.

Seon-Jong Kim obtained his PhD in electronics from Kyungpook National University, Korea in august 1996. He joined Suncheon 1 College in February 1995 and Pusan National University in March 1997, and he is working as professor at department of IT engineering in Pusan National University currently. He had experienced a visiting scholar for one year in Georgia State University in 2008. He is teaching the subjects to digital system and design, such as logical design, microprocessors, and signal systems. His research interest includes signal and image processing, machine vision, AR/VR in general.