2009 Symposium on Bio-inspired Learning and Intelligent Systems for Security

Recognition of Dynamic Texture Patterns Using CHLAC Features

Takumi Kobayashi† Tetsuya Higuchi† Tsuneharu Miyajima‡ Nobuyuki Otsu†

†National Institute of Advanced Industrial Science and Technology, 1-1-1 Umezono, Tsukuba, Japan
‡Saga Prefectural Livestock Experiment Station, Takeo, Saga, Japan

Abstract

In this paper, we propose a statistical scheme for recognizing three-dimensional textures shown in motion images, which we call dynamic textures. The texture characteristics emerges in the distinct movement in the motion images, and the dynamic cues would be useful especially for recognizing ambiguous texture patterns in noisy images. We apply cubic higher-order auto-correlation (CHLAC) to extract features both of the textures and their movements, and then simply multiple regression analysis (MRA) to evaluate (recognize) the texture. In the experiment for estimating quality of beef meat by using ultrasound motion images, the proposed method exhibits the favorable performances which are close to ground truth given by the experts.

1. Introduction

Texture classification has been an important research topic in image processing and it provides important cues for many applications, such as object recognition and image retrieval. Many researchers have proposed successful methods [1] to classify the textures in a still image, we call static texture. On the other hand, the dynamic textures that appear in motion images have rarely been dealt with. The dynamic cues in such textures are useful to recognize them especially when the texture patterns can not be obviously recognized in an static image due to noises, such as in ultrasound images. In fact, humans can perceive the distinct patterns by the movement, as stated by psychologists [2]. In addition, the traditional approaches to extract texture features, e.g., derivative filter responses [1], gray-level cooccurrence matrix [4] and wavelets [5], are not so applicable to the noisy images.

In this paper, we propose a statistical scheme for recognizing the dynamic textures in motion images. For extracting features of the texture patterns, we employ the method of cubic higher-order auto-correlation (CHLAC) [8] which has been mainly applied to motion recognition. CHLAC can naturally deal with the (three-dimensional) dynamic texture patterns whereas HLAC [9] that extracts features from static images is applied to the classification of static textures. The method of CHLAC extracts the features of the dynamic textures by local auto-correlations of pixel values. After extracting the texture features, we can estimate the class label or the quantitative scores associated with the dynamic texture by applying, such as multivariate analysis [6] and SVM.

Recent years, it has been shown that ultrasound images are useful to estimate carcass traits in the live animal, that is, there are somewhat high correlations between ultrasound images and carcass traits [3]. The estimation of carcass traits, however, is manually performed by experts, and it is desirable to automate the estimation from the ultrasound images. Therefore, we apply the proposed method to automatically estimate the quality of beef meat, i.e., beef marbling score (BMS), by using ultrasound motion images. In the experiment, the proposed method produces the favorable performances compared to the ground truth given by the experts, although the ultrasound images contain a large amount of noise complicating the recognition in static images.

2. Proposed Method

The proposed method for recognizing dynamic textures is composed of 1) CHLAC [8] to extract features from three-way dynamic texture patterns in motion images and 2) MRA [6] to estimate the quantitative scores from the extracted features.

2.1. Feature Extraction

The dynamic texture appears in a motion image sequence, not in a still image. We describe the characteristics of the texture in the sequence by utilizing assembles of feature vectors rather than a single feature vector. Thereby, the dynamic textures that change with times are effectively characterized by each feature vector at each time. The method of CHLAC [8] can be applied to extract features of the dynamic texture in the motion images.

978-0-7695-3754-2/09 $25.00 © 2009 IEEE
DOI 10.1109/BLISS.2009.31

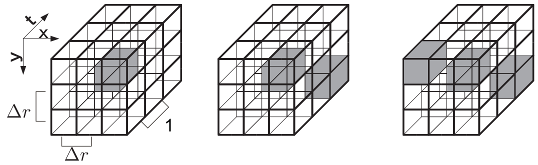

Figure 1. CHLAC mask patterns. Δr indicates the spatial interval of correlations.

CHLAC is based on the auto-correlations of pixel values within the motion image in X, Y and T. The auto-correlations are actually calculated by using 279 mask patterns (e.g., Fig. 1): the pixel values indicated in the mask patterns are multiplied and summed up over the whole image region. In this study, from a single sequence, we sample substantial subsequences of which time duration is denoted by T. CHLAC feature (279-dimensional vector) is extracted from each subsequence and we finally obtain a large number of feature vectors from the single sequence.

While derivative filter responses, such as Gaussian derivatives and Gabor filters, are often employed as texture features in the traditional methods for the static texture classification, we extract the correlation-based features of the dynamic texture. The dynamic texture shown in motion images are noisy, especially in ultrasound images (Fig. 2), and thus the responses of the derivative filters are noise-sensitive and unstable. On the other hand, CHLAC is based on auto-correlations which are known to be robust to noises. It should be noted that optical flows are still less feasible for such noisy motion images.

2.2. Recognition

Next, we estimate the quantitative score associated with the dynamic texture. Given scores for respective sequences, we apply multiple regression analysis (MRA) which determines the optimal linear coefficients to estimate the scores from feature vectors. Let $\boldsymbol{x}_j^{(i)}$ be the j-th feature vector extracted from the i-th sequence and y_i be the score for the i-th sequence. The MRA results in the following optimization problem:

$$\min_{\boldsymbol{a},b} \sum_i \sum_j ||y_i - \boldsymbol{a}'\boldsymbol{x}_j^{(i)} - b||^2$$
$$\Rightarrow \boldsymbol{a} = (\bar{\boldsymbol{X}}\bar{\boldsymbol{X}}')^{-1}\bar{\boldsymbol{X}}\bar{\boldsymbol{y}}, \qquad b = \bar{y} - \bar{\boldsymbol{x}}'(\bar{\boldsymbol{X}}\bar{\boldsymbol{X}}')^{-1}\bar{\boldsymbol{X}}\bar{\boldsymbol{y}}$$

where

$$\bar{\boldsymbol{X}} = [\boldsymbol{x}_1^{(1)} - \bar{\boldsymbol{x}}, \cdots, \boldsymbol{x}_{n_N}^{(N)} - \bar{\boldsymbol{x}}],$$
$$\bar{\boldsymbol{y}} = [\underbrace{y_1 - \bar{y}, ..., y_i - \bar{y}}_{n_1}, \cdots, \underbrace{y_N - \bar{y}, ..., y_N - \bar{y}}_{n_N}]',$$
$$\bar{\boldsymbol{x}} = E_{ij}(\boldsymbol{x}_j^{(i)}), \qquad \bar{y} = E_i(y_i),$$

Figure 2. Example of ultrasound image. The target region is indicated by red curves.

$n_i (i = 1, \cdots, N)$ is the number of features (subsequences) extracted from the i-th sequence, and N is the number of sequences. The score for the j-th feature (subsequence) in the input sequence is estimated by

$$\hat{y_j} = \boldsymbol{a}'\boldsymbol{x}_j + b,$$

and then the final score for the whole sequence is predicted by averaging above estimation results:

$$\hat{y} = E_j(\hat{y_j}) = \frac{1}{n}\sum_{j=1}^{n} \hat{y_j}.$$

In this study, the MRA is applied on the assumption that the target is quantitative score value. MRA with CHLAC/HLAC features has produced favorable performance for multiple person identification [7] and counting objects [6]. Note that, in the case that the target is qualitative class label, we can employ classifiers, such as SVM.

3. Experimental Result

We apply the proposed method to estimate quality of beef meat, i.e., beef marbling score (BMS), from ultrasound motion images as shown in Fig. 2. The ultrasound images would enable us to estimate BMS without slaughtering cattle, which is quite useful for the breeders. It is, however, difficult to identify the textures in ultrasound images because the texture characteristics are ambiguous due to a large amount of image noses.

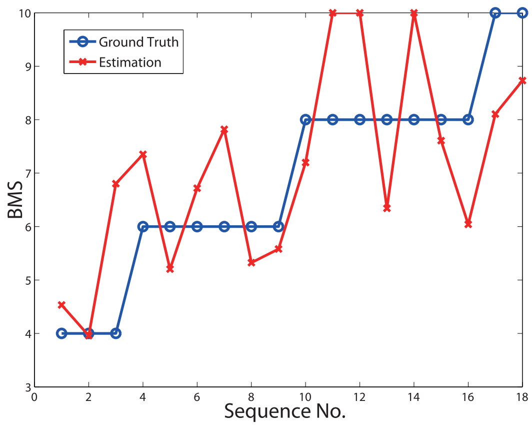

Figure 3. Performance results

We have 18 sequences for various BMS: three sequences for BMS-4, six sequences for BMS-6, seven sequences for BMS-8, two sequences for BMS-10. CHLAC with various correlation intervals ($\Delta r \in \{1, \cdots, 8\}$) is extracted from subsequence whose time durations are $T \in \{10, 20, 30, 40, 50\}$. Thus, the dimensionality of the feature vector is $279 \times 8 = 2232$. Note that the CHLAC features are extracted within the target image region, as shown in Fig. 2. In this experiment, the BMS is limited up to 10 and therefore the upper limit of the estimation $\hat{y}$ is also constrained to 10: $\hat{y} \leftarrow \min(\hat{y}, 10)$.

The performance is evaluated based on leave-one-sequence-out method for the 18 sequences. The mean absolute error between estimated and ground truth BMS is *1.29*. The results of the estimations for each sequence are shown in Fig. 3. Although the ultrasound images contain a great amount of noises which makes the estimation quite difficult, the proposed method extracts the characteristics of dynamic textures related to BMS and produces the favorable performances.

4. Conclusion and Future Work

We have proposed a statistical scheme for recognizing dynamic texture patterns shown in motion image sequences. In the proposed scheme, the characteristics of the dynamic textures are efficiently extracted by CHLAC, and the target score is estimated from the extracted feature by multiple regression analysis. CHLAC is based on the auto-correlation of the dynamic texture patterns and is robust to noises contained in the motion images. Even in the preliminary experiment of estimation of beef meat quality (BMS) by using ultrasound motion images, the proposed method produced the favorable performances, in spite that the ultrasound images contain a large amount of noises. The result shows that CHLAC can extract sufficient texture features from the ultrasound motion images.

The obtained results are sufficient but a little further improvement is required for industrial applications. The future work for estimating BMS includes applying more sophisticated regression method, such as kernel MRA and support vector regression, which would produce more favorable estimation results.

References

[1] M. Crosier and L. Griffin. Texture classification with a dictionary of basic image features. In *IEEE Conference on Computer Vision and Pattern Recognition*, 2008.

[2] J. Gibson, editor. *The Ecological Approach to Visual Perception*. Houghton Mifflin, 1979.

[3] S. P. Greiner. Ultrasound applications for the beef industry. *Live Stock Update*, April 2003.

[4] R. M. Haralick, K. Shanmugam, and I. Dinstein. Textural features for image classification. *IEEE Transaction on Systems, Man, and Cybernetics*, SMC-3(6):610–621, 1973.

[5] N.-D. Kim, V. Amin, D. Wilson, G. Rouse, and S. Udpa. Ultrasound image texture analysis for characterizing intramuscular fat content of live beef cattle. *Ultrasonic Imaging*, 20:191–205, 1998.

[6] T. Kobayashi, T. Hosaka, S. Mimura, T. Hayashi, and N. Otsu. Hlac approach to automatic object counting. In *The 2008 ECSIS Symposium on Bio-inspired, Learning, and Intelligent Systems for Security (BLISS)*, pages 40–45, 2008.

[7] T. Kobayashi and N. Otsu. Action and simultaneous multiple-person identification using cubic higher-order local auto-correlation. In *International Conference on Pattern Recognition*, pages 741–744, 2004.

[8] T. Kobayashi and N. Otsu. Three-way auto correlation approach to motion recognition. In *Pattern Recognition Letters*, volume 30, pages 212–221, 2009.

[9] N. Otsu and T. Kurita. A new scheme for practical flexible and intelligent vision systems. In *IAPR Workshop on Computer Vision*, 1988.