Brief Descriptions of Visual Features for Baseline TRECVID Concept Detectors

Akira Yanagawa, Winston Hsu and Shih-Fu Chang
{akira, winston, sfchang@ee.columbia.edu}

1. INTRODUCTION

We describe the visual features used in the implementations of baseline concept detectors for our participation in NIST TRECVID 2006 high-level feature detection. Such features are part of the rich feature set that has been shown to be effective in previous TRECVID experiments, first in [1] then repeated in [2] and [3]. Original algorithms for extracting each feature may be found in references [4-6].

Publications of research using the feature data included here may acknowledge Akira Yanagawa and Digital Video/Multimedia Lab of Columbia University, and cite this technical report. References to original algorithms, when needed, should be made to those listed in the reference section.

2. Edge Direction Histogram

An Edge Direction Histogram denotes the direction of edge, literally [6]. The number of the bins is 73: 72 for edge direction quantized at 5 degrees and one bin is for non-edge points. To detect edge points, we used a Canny filter, and to get the direction, we calculated the gradient of each edge point by utilizing a Sobel operator. This histogram is normalized by the number of all pixels to cancel the effect of the image size.

3. Gabor Filter

Gabor filter banks are found to be useful for analyzing texture for image retrieval [4]. In [4], they derived a two-dimensional Gabor function to get the self-similar dictionary adjusted by two parameters: scale and orientation. To denote the texture in an image, we used the mean and standard deviations of the output of a Gabor filter by using the combinations of four scales and six orientations.
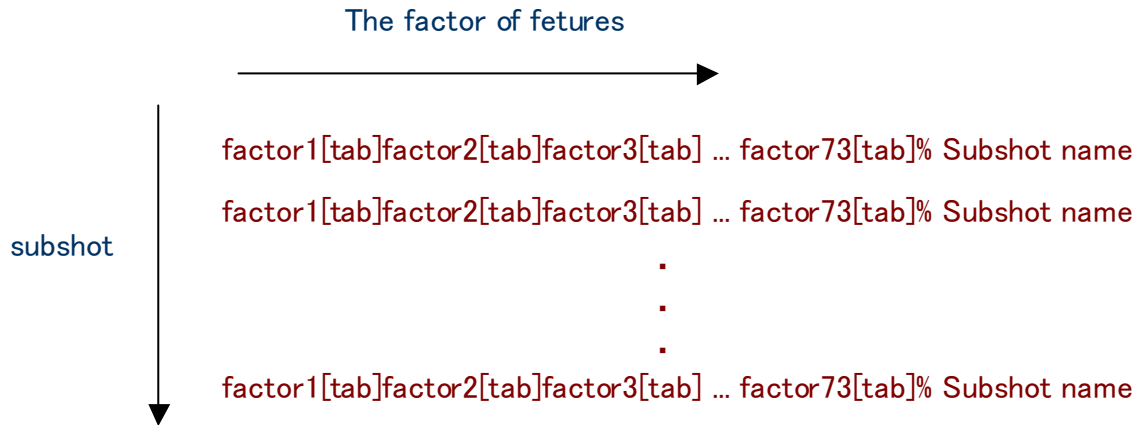
4. Grid Color Moment

Color histograms are useful in describing the color distribution in an image. However, since there is no local relationship between information in color histograms, it cannot always differentiate between objects. To compensate for this, we utilized grid color moment, in addition to color histograms.

Color moment denotes the color distribution by using mean, standard deviation, and the third root of the skewness of each color channel[5]. If the value of the ith color channel at the (x,y) image pixel is p_{xy}^i, and the width and height of grid are M and N respectively then the color moments are:
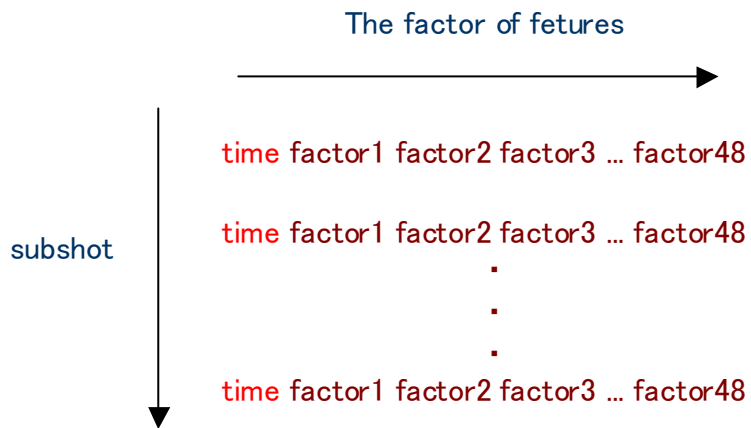
$$E^i = \frac{1}{MN}\sum_{y=y_s}^{y_e}\sum_{x=x_s}^{x_e} p_{x,y}^i \quad \sigma^i = \left[\frac{1}{MN}\sum_{y=y_s}^{y_e}\sum_{x=x_s}^{x_e}\left(p_{x,y}^i - E^i\right)^2\right]^{\frac{1}{2}} \quad s^i = \left[\frac{1}{MN}\sum_{y=y_s}^{y_e}\sum_{x=x_s}^{x_e}\left(p_{x,y}^i - E^i\right)^3\right]^{\frac{1}{3}}$$

Where $x_s, x_e, y_s \ and \ y_e$ are the start and end x coordinate and the start and end y coordinate, respectively. For color space, we used LUV, which is suitable for image recognition. In addition, we divided images into 5x5 grids.
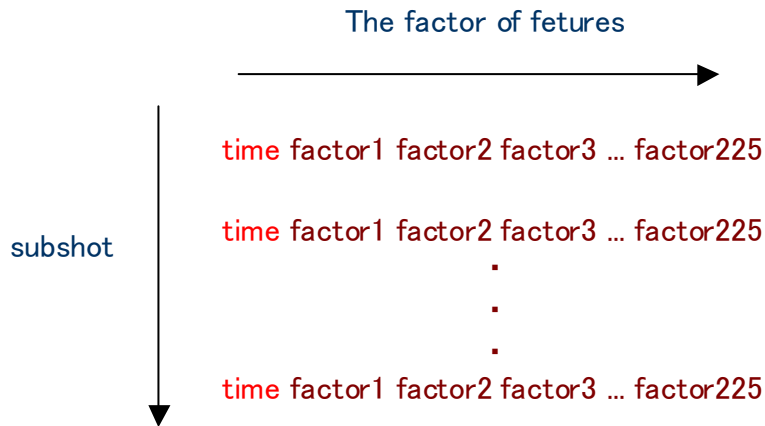
5. File Format: Edge Direction Histogram (73 dimensions)

The factor of fetures

subshot

factor1[tab]factor2[tab]factor3[tab] … factor73[tab]% Subshot name

factor1[tab]factor2[tab]factor3[tab] … factor73[tab]% Subshot name

.
.
.

factor1[tab]factor2[tab]factor3[tab] … factor73[tab]% Subshot name

6. File Format: Gabor (48 dimensions)

The factor of fetures

subshot

time factor1 factor2 factor3 … factor48

time factor1 factor2 factor3 … factor48

.
.
.

time factor1 factor2 factor3 … factor48

7. File Format: Grid Color Moments (225 dimensions)

The factor of fetures
→

subshot ↓

time factor1 factor2 factor3 … factor225

time factor1 factor2 factor3 … factor225

.
.
.

time factor1 factor2 factor3 … factor225

8. References

[1] Arnon Amir, Janne O Argillander, Marco Berg, Shih-Fu Chang, Martin Franz, Winston Hsu, Giridharan Iyengar, John R Kender, Lyndon Kennedy, Ching-Yung Lin, Milind Naphade, Apostol (Paul) Natsev, John R. Smith, Jelena Tesic, Gang Wu, Rong Yan, Donqing Zhang, "IBM Research TRECVID-2004 Video Retrieval System," In NIST TRECVID 2004 Workshop, Gaithersburg, MD, 2004.

[2] Shih-Fu Chang, Winston Hsu, Lyndon Kennedy, Lexing Xie, Akira Yanagawa, Eric Zavesky, Dong-Qing Zhang, "Columbia University TRECVID-2005 Video Search and High-Level Feature Extraction," in NIST TRECVID workshop, Gaithersburg, MD, Nov. 2005.

[3] A. Amir, J. Argillandery, M. Campbell, A. Haubold, G. Iyengar, S. Ebadollahi, F. Kang, M. R. Naphade, A. P. Natsev, J. R. Smith, J. Tesic, and T. Volkmer., "IBM Research TRECVID-2005 Video Retrieval System," In NIST TRECVID 2005 Workshop, Gaithersburg, MD, Nov 2005.

[4] B. S. Manjunath, W. -Y. Ma, "Texture Features for Browsing and Retrieval of Image Data", IEEE Trans on PAMI, Vol. 18, No. 8, pp. 837-842, 1996.

[5] M. A. Stricker and M. Orengo, "Similarity of color images," in Proceedings of SPIE, Storage and Retrieval for Image and Video Databases III, vol. 2420, March 1995, pp. 381-392.

[6] A. Vailaya, A. Jain, and HJ Zhang, "On Image Classification: City Images vs. Landscapes," Pattern Recognition Journal, Pattern Recognition, Vol. 31, No. 12, pp. 1921-1935, 1998