

# 영상의 질감 정보를 이용한 검색에서 회전 불변 유사도 측정에 대한 연구

## Study of Rotation invariant similarity measurement for content based image indexing using Texture

강호경, 유기원, 최정윤, 노용만

한국정보통신 대학원 대학교, 영상/비디오 연구실

### 요 약

본 논문에서는 MPEG-7 질감 기술자에 기반한 영상의 내용기반 검색 시스템에서 유사도 검색 알고리즘에 대하여 다룬다. 제안하는 알고리즘은 영상의 회전에 대하여 변화하지 않는 특성을 지닌다. 실제적으로 영상이나 비디오에서 질감 정보를 이용하여 검색을 수행하고자 하는 영상에 대하여 회전 불변 알고리즘은 반드시 필요하다. 본 논문에서 제한된 알고리즘을 이용하여 질감 영상 회전 불변 유사도 측정에 따른 검색을 수행하였다. 본 논문의 실험은 회전 불변 성능을 측정하는 MPEG-7 질감 데이터베이스에서 검색 율을 측정하였다. 실험결과 본 영상 회전 불변 질감 기술자 표현 방법은 우수한 검색 성능과 빠른 특징 추출 능력을 보였다.

### I. 서 론

최근의 저장 매체 및 네트워크 통신의 발달과 더불어 비주얼 멀티미디어 데이터베이스의 크기는 기하급수적으로 증가하고 있다. 이러한 방대한 양의 비주얼 멀티미디어 데이터의 효율적 관리, 검색 및 인식의 필요성과 더불어 현재 진행 중인 MPEG-7 과 관련한 멀티미디어 데이터의 관리 및 검색 기술들의 표준화

과정에서 영상의 내용 특징에 기반한 다양한 검색 알고리즘들이 개발되고 있다. 텍스처 정보는 색, 윤곽, 움직임 정보와 함께 영상의 내용 정보를 기술하는 가장 기본이 되는 특징의 하나로, 이러한 텍스처 특징을 추출하는 기술은 영상의 내용 정보를 기술하는 기반 기술이 된다.

영상의 질감 정보는 색, 윤곽, 움직임 등과 함께 영상의 내용 정보를 기술하는 중요한 시

각적인 특징의 하나로, 의료, 우주 항공, 영상 편집 및 대기 영상 등의 영상 분석, 객체 인식, 표면 분석 등과 관련하여 최근 30여 년 간 폭넓게 연구되어오고 있다. 일상에서 자주 접하는 질감 정보로는 옷감의 무늬, 벽, 타일 등의 균질 또는 비 균질의 패턴을 예로 들 수 있으며, 이러한 질감은 영상 화소간의 주기성 또는 반주기성으로 설명된다. 이 질감정보는 영상의 구조나, 방향성, 거친 정도 등의 특성을 나타내고, 영상 또는 비디오 데이터를 내용 기반 색인(indexing)하거나 요약하는데 있어서 중요한 특징으로 사용된다.

본 논문에서는 MPEG-7 비주얼 부분의 위원회 작업 초안(Working Draft)에 채택된 HVS 특성에 의한 질감 기술자의 회전 불변으로 검색할 수 있는 유사도 계산방법에 대한 제안한다

본 논문의 구성은 다음과 같다. 2장 이론에서 MPEG-7 질감 특징 추출 알고리즘에 대한 소개를 하고 회전 불변 질감 검색 알고리즘에 대하여 제안한다. 3 장에서는 실험 및 결과를 보인다. 그리고 마지막으로 결론을 내린다.

## II. 이론

### 1. 질감 특징을 위한 HVS 구조

위의 HVS 특성을 따라서 본 논문에서는 질감 특징 추출에 사용되는 채널을 각도 방향과 방사 방향으로 다음 그림과 같이 분할함으로써 얻었다. 그림 1은 방사방향에서 영역을 2의 배수로 채널의 크기가 증가 하면서 5 개로 분할하고 각도 방향으로는 30 도 마다 분할하면서 주파수 공간을 나누었다.

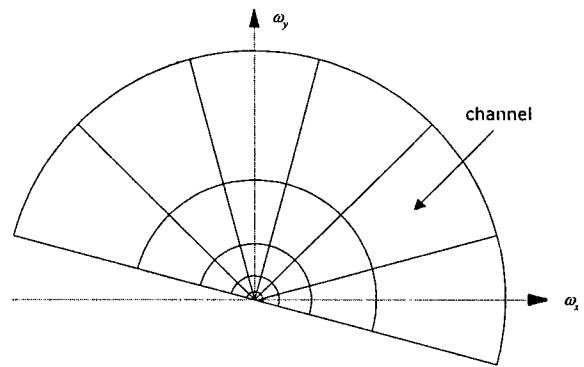


그림 1. HVS 채널에 의한 주파수 공간의 분해

### 2. 질감 특징

질감 특징은 HVS 로 나누어진 각 채널에서 에너지 벡터와 에너지 편차로 계산되어진다. 에너지를 사용한 질감 특징 벡터는 다음과 같이 기술되고 각 채널의 주파수 공간에서의 값을  $P_{\theta}(\omega)$ 라고 하면

$$p_{mn} = C_{mn} \sum_{\omega} \sum_{\theta} [P_{\theta}(\omega)]^2 \quad (1)$$

$$e_{mn} = \log(1 + p_{mn}) \quad (2)$$

그리고 에너지 편차를 이용한 질감 특징 벡터는 다음과 같다.

$$q_{mn} = \sqrt{D_{mn} \sum_{\omega} \sum_{\theta} [(P_{\theta}(\omega))^2 - p_{mn}]^2} \quad (3)$$

$$d_{mn} = \log(1 + q_{mn}) \quad (4)$$

여기서  $C_{MN}$  와  $D_{MN}$  는 정규화 값을 구하기 위한 상수이다. 이러한 값과 한 영상의 평균값과 표준 편차를 사용하여 질감 기술자는 다음의 수식 5로 표현된다.

$$Descriptor_{texture} = \begin{bmatrix} dc & std & e_{00} & e_{01} & \dots \\ e_{45} & d_{00} & d_{01} & \dots & d_{45} \end{bmatrix} \quad (5)$$

### 3. 회전 불변 유사도 측정

검색하고자 하는 이미지와 비슷한 이미지를 찾기 위해선 분류가 필요하다. 이 논문에서는 이를 위한 단계가 다음과 같다. 극좌표에서 주파수 데이터가 얻은 다음에 질감 특징을 기술하기 위해 앞에서 설명한 에너지와 편차가 계산된다. 데이터 베이스와의 유사도를 측정하기 위해서 검색하고자 하는 이미지의 특징을  $TD_{query}$  라 하고 데이터 베이스의 질감 특징을  $TD_{Database}$  라고 할 때 유사도는 다음과 같이 계산되어 질 수 있다

$$d(TD_{query}, TD_{Database}) = distance(TD_{query}, TD_{Database}) \\ = \sum_k \left| \frac{TD_{query}(k) - TD_{Database}(k)}{\alpha(k)} \right| \quad (6)$$

여기서  $\alpha(k)$  는  $TD_{Database}(k)$  의 표준 편차이다.

질감 특징 기술자는 2 차원 푸리에 영역이므로 입력 영상의 회전에 대해서 채널들이 각도 방향으로 회전한 만큼의 각도로 이동하는 특성이 있다. 따라서 회전불변 검색을 위한 거리 측정에서 회전 각도를 고려하면 회전 불변의 검색을 수행할 수 있다.

거리 함수를  $\phi$  에 의한 함수로 놓으면 식 7 와 같이 된다.

$$d(i, j, m\phi) = distance(TD_{query}(k)_{m\phi}, TD_{Database}(k)) \quad (7)$$

각도  $\phi$  는  $180/L$  값으로 단위 각도가 된다. m

은 1 에서 L 까지 정수이다. 회전 불변 질감 특징 유사도 검색은 식 8 에 의한 거리 함수에 의하여 측정된다.

$$d(i, j) = \min\{d(i, j, m\phi) | m = 1 \text{ to } L\} \quad (8)$$

### 4. 계층적인 회전 불변 유사도 측정

식 8 에 따르면 회전 불변 유사도를 측정을 위하여 원하는 회전 각도에 따라서 모두 거리 측정을 수행하여야 한다. 이러한 방법은 만약 5 단위로 유사도 측정을 수행하는 경우 한 영상에 대하여 36 번의 거리 측정을 수행해야 하는 단점이 존재한다.

따라서 회전 불변 유사도 측정시 측정 횟수를 단축하기 위하여 다음과 같은 계층적인 유사도 검색을 수행할 수 있다.

예로 30 단위로 6 번의 검색을 수행하여 가장 가까운 부분을 선택하고 선택된 부분으로부터 10 단위로 4 번의 검색을 수행한 뒤 5 단위로 2 번의 유사도 측정을 수행하여 3 단계로 회전 불변 유사도 측정을 수행한다.

단계 1: 다음 식을 만족하는  $\theta_{ref}^1 = m\phi$  을 구한다.

$$\min\{d(i, j, m\phi) | m = 1 \text{ to } 6, \phi = 30^\circ\} \quad (9)$$

단계 2: 다음 식을 만족하는  $\theta_{ref}^2 = \theta_{ref}^1 + \Delta\theta$  를 구한다.

$$\min\{d(i, j, \theta_{ref}^1 + \Delta\theta) | \Delta\theta = \pm 10^\circ, 0^\circ, \pm 20^\circ\} \quad (10)$$

단계 3: 최종적인 거리를 측정한다.

$$d(i, j) = \min\{d(i, j, \theta_{ref}^2 + \Delta\theta) | \Delta\theta = 0^\circ, \pm 5^\circ\} \quad (11)$$

여기서  $\theta_{ref}$  는 각 레벨에서 최소거리를 구할 때 기준이 되는 각도를 의미하고  $\Delta\theta$  는 식 9, 10 에서 보이듯이 2 단계에서는 10 단위로 3 단계에서는 5 단위로 변화를 하면서 거리를 측정하게 된다.

그림 2 은 3 단계의 계층적인 회전 불변 유사도 측정의 예를 보인다. 이 방법에 의하여 유사도 측정을 12 번으로 줄일 수 있다.

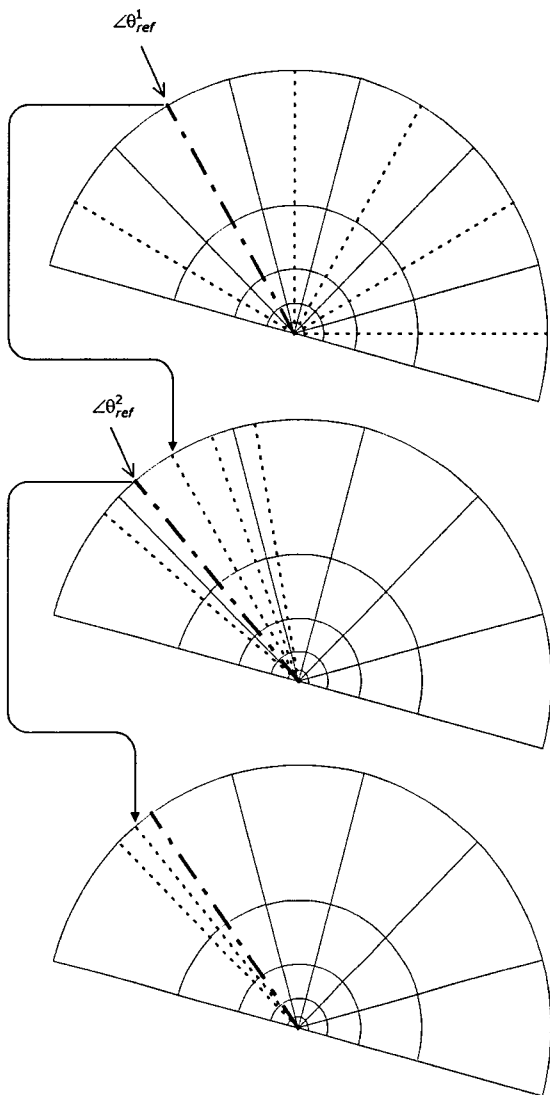


그림 2. 레벨 계층적인 유사도 측정

#### IV. 실험 및 결과

회전 불변 특징 알고리즘의 성능을 평가하는

T3 데이터는 MPEG-7 에서 제안된 DB 이다. 이때 회전 불변에서 선형 보간에 의한 임의의 각에서의 특징 벡터의 유도 시에  $\theta$  는  $5^\circ$  로 설정하였다.  $5^\circ$  간격에서 유도되는 특징 벡터의 수는  $36(=180/5)$ 개가 되며, 질의 영상  $i$  와 데이터베이스내의 영상간의 거리 함수는 36 번 계산된다.

제안 알고리즘에서는 이때 소모되는 계산량의 감소를 위해 계층적인 방법(HDM : hierarchical distance matching)을 사용하였다.

제안 알고리즘의 실험 과정에서는  $\theta_1 = 30^\circ$ ,  $\theta_2 = 10^\circ$ ,  $\theta_3 = 5^\circ$  를 사용하며, 이때 사용되는 거리 함수의 계산은 단계 1)에서 6 번( $=180/6$ ), 단계 2), 단계 2)에서 4 번( $=40/10$ ), 단계 3)에서 2 번( $=10/5$ )으로 모두 12 번의 거리 함수 계산 과정이 사용된다. 이것은 계층적인 방법을 사용하지 않았을 때에 사용되는 36 번의 계산량을 1/3 로 감소한 것이다. 이러한 과정을 그림 2 에 나타내었다.

표 1 은 T3 데이터를 이용하여 회전 불변과 관련하여 사용되는 다양한 유사도 측정 과정에 따른 검색 결과들을 비교한 것이다.

표 1. 회전 불변 알고리즘에 따른 검색 결과

사용 방법		검색율 (%)	시간 (msec)
계층적 방법 (HDM)	단계 1 (6 번)	92.73	6.65
	단계 2 (10 번)	92.48	11.1
	단계 3 (12 번)	90.06	13.5
Full matching (36 번)		92.36	39.9
기준 축 이용		92.85	1.1

표 1 에서 보이듯이, 계층적인 방법을 사용하였을 때의 검색 결과는 전체 거리 함수의 계

산을 모두 수행하였을 때 보다 단계 2)까지 보다 나은 성능을 보이며, 검색 과정에서 시간과 검색 성능과의 절충 과정에서 사용자의 요구 사항을 효과적으로 반영할 수 있다는 장점이 있다.

## 결 론

본 논문에서는 MPEG-7 의 질감 기술자의 회전 불변 알고리즘에 대하여 제안하였다. 현 알고리즘은 기존의 특징 추출 알고리즘의 변화 없이 바로 이용이 가능하다. 다른 회전 불변 알고리즘에 대하여 대한 실험은 기존의 질감 성분 추출하는 알고리즘의 변화를 요하고 또한 회전 불변 DB 에만 좋은 성능을 지니고 다른 DB 에서는 많이 성능이 저하되는 것을 확인하였다. 이 제안한 방법은 허용가능한 시간에 최적의 성능을 내고 있다.

## 참고 문헌

- [1] MPEG-7 Visual part of eXperimentation Model Version 6.0, *ISO/MPEG*, n3398, Geneva, June, 2000
- [2] MPEG-7 Visual Working Draft 3.0, *ISO/MPEG*, n3399, Geneva, June, 2000
- [3] O.D.Faugeras, W.K.Partt, Decorrelation methods of texture feature extraction, *IEEE Trans. Pattern Anal. and Machine Intell.*, vol.2, pp. 323-332, July 1980
- [4] B.S.Manjunath, W.Y.Ma, Texture Features for Browsing and Retrieval of Image Data, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, No. 8, August 1996
- [5] J.R.Ohm, F.Bunjamin, Descriptor for texture in wavelet domain, *ISO/MPEG*, P566, Lancaster, 1999
- [6] A.Saadane, H.Senane, D.Barba, On the Design of Psychovisual Quantizers for a Visual Subband Image Coding, *SPIE*, Vol. 2308, pp1446, 1994
- [7] J.G.Daugman, High Confidence Visual Recongnition of Persons by a Test of Statistical Independence, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol.15, no.11, pp.1148-1161, November, 1993
- [8] C.J.Lambrecht, A Working Spatio - Temporal Model of Human Vision System for Image Restoration and Quality Assessment Applications, *IEEE International Conference on ASSP*, New York, NY, USA, p.2291-4 vol. 4., 1996
- [9] Y.M.Ro, S.Y.Kim, K.W.Yoo, M.C.Kim, J.W.Kim, Texture descriptor using atoms for matching pursuit, *ISO/MPEG*, P612, Lancaster 1999
- [10] Y.M.Ro, K.W.Yoo, M.C.Kim, J.W.Kim, Texture Description using Radon transform, *ISO/MPEG*, m4703, Vancouver, 1999
- [11] Y.M.Ro, K.W.Yoo, M.C.Kim, J.W.Kim, Texture description using Radon transform and experimental results on CT-5 core experiment using atoms for matching pursuit, *ISO/MPEG*, m5152, Melbourne 1999
- [12] Y.M.Ro, K.W.Yoo, M.C.Kim, J.W.Kim, B.S.Manjunath, D.G.Sim, H.K.Kim, J.R.Ohm, An unified texture descriptor *ISO/MPEG*, m5490, Maui 1999