

# DFT를 이용한 Face Detection

## DFT integration for Face Detection

한석민\*, 최진영\*\*  
(Seok Min Han, Jin Young Choi)

**Abstract** - In this work, we suggest another method to localize DFT in spatial domain. This enables DFT algorithm to be used for local pattern matching. Once calculated, it costs same load to calculate localized DFT regardless of the size or the position of local region in spatial domain. We applied this method to face detection problem and got the results which prove the utility of our method.

**Key Words** :Face Detection, DFT, Shift Invariance, Pattern Matching

### 1. 서론

DFT(Discrete Fourier Transform)은 순차적인 신호를 주파수 신호로 변환하는 방식으로써, 영상신호처리, 음성신호처리등의 분야에서, 신호의 압축이나 차원감소(Dimension Reduction)을 위해 널리 쓰이고 있는 방식이다. DFT를 통해 순차신호를 주파수 영역으로 전환하면, 대부분의 에너지가 저주파 영역(Low Frequency)에 위치하기 때문에, 이를 이용하여 영상 압축, 음성압축등의 신호압축에 DFT를 이용할 수 있는데, JPEG등 신호 압축에 널리 이용되는 DCT는 바로 이 DFT를 통해 얻어진 신호변환 방식이다. DFT는 에너지 압축의 성질 뿐 아니라, 신호가 시간영역에서 변이(Time shift)되더라도 DFT로 변환된 신호의 크기(magnitude)는 변하지 않는다는 특징이 있다. 이 성질은 동전의 양면과 같이 장단점을 동시에 지니고 있다. 우선 장점은, 변환된 신호의 크기(magnitude)가 시간영역에서의 변이에 영향을 받지 않는다는 점이다. 단점은, 시간영역에서의 변이에 영향을 받지 않기 때문에, 어떤 주파수가 어떤 시간 영역에서 발생한 것인지 알 수 없다는 점이다. 이 때문에 DFT는 패턴 매칭(Pattern Matching)에서 사용되기 곤란한 측면이 있다. 하지만, WFT(Window Fourier Transform)등과 같이, 푸리에 변환(Fourier Transform)의 영역을 적절하게 제한한다면, 주파수 영역에서도 원하는 신호의 존재 여부를 가늠할 수 있게 된다. 본 논문에서는 WFT와 같이 실제 신호의 영역을 제한함으로써, 주파수 영역에서도 원하는 신호의 존재 여부를 확인할 수 있는 DFT 계산 알고리즘을 제시하고, 이 알고리즘을

Face Detection에 적용하도록 한다. 이를 통해 공간적으로 제한된 영역의 DFT가 패턴 매칭에 유용하게 이용될 수 있음을 보이도록 한다.

### 2. 적분형 DFT

2D-DFT는 주어진 영상의 모든 영역에 대해 DFT를 수행한다. 따라서, 특정한 영역별 DFT를 구하려면 일반적인 DFT를 변형하여, 영역별 DFT를 효율적으로 계산할 수 형태를 갖추도록 해야한다. 이를 위하여 적분형 DFT를 고안하여, 영상 안의 어떤 영역이라도(Rectangle 형 영역) 4개의 좌표를 알고 있으면 곧바로 그 영역의 DFT를 구할 수 있게 하였다. 또한 구해진 DFT에서 패턴 매칭 연산을 수행, 원하는 패턴과 유사한 영역을 검출하도록 한다.

#### 2.1 적분형 DFT의 생성

크기가  $N \times N$  인 영상 이미지의 DFT변환식은 다음과 같다.

$$F(k, l) = \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} I(m, n) \exp^{-j\frac{2\pi}{N} km} \exp^{-j\frac{2\pi}{N} ln} \quad (1)$$

(1)식은 spatial frequency  $(k, l)$ 에 대한, 영상이미지  $I(m, n)$  전체에 대한 DFT 변환식이다. ( $N$ 은 주기)  $F(l, k)$ 도  $I(m, n)$ 과 마찬가지로  $N \times N$  의 크기를 갖게 된다.

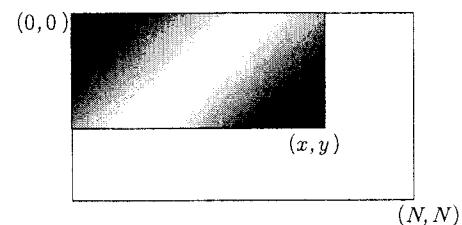


그림 1 (x,y)까지의 DFT 변환

#### 저자 소개

\* : 서울대학교 전기컴퓨터공학부 博士課程

\*\* : 서울대학교 전기컴퓨터공학부 教授

(1)식을 변형하여, 그림 2 와 같이  $(0,0) \sim (x,y)$  까지의 픽셀들에 대해서만 DFT변환을 수행하도록 한다. 그럼 다음과 같은 식을 갖는다.

$$F(x,y,k,l) = \sum_{m=0}^{x-1} \sum_{n=0}^{y-1} I(m,n) \exp^{-\frac{j2\pi}{N} km} \exp^{-\frac{j2\pi}{N} ln} \quad (2)$$

(2)식을  $0 \leq x, y \leq N-1$  인  $x, y$ 에 대해 계산한다. 이 계산을 빠르게 하기 위해, FFT(Fast Fourier Transform)을 이용할 수 있다.

$$F_x(x,n,k) = \sum_{m=0}^{x-1} I(m,n) \exp^{-\frac{j2\pi}{N} km} \quad (3)$$

(3)의 식은 1D-FFT를 이용하여 빠르게 계산할 수 있다.

$$F(x,y,k,l) = \sum_{n=0}^{y-1} F_x(x,n,k) \exp^{-\frac{j2\pi}{N} ln} \quad (4)$$

즉, (2)의 식은 (3)을 이용하여 (4)의 식으로 정리된다. 일반적인 DFT를 이용한다면 각 좌표에 대해  $O(N^4)$ 의 계산이 필요하겠지만( $N \times N$  이미지에 대한 연산량은  $O(N^4)$ ), (3)식을 1D-FFT를 이용한다면, 1D-FFT의 연산량이  $O(N \log_2 N)$  이므로[1] (4)식은  $O(N^3 \log_2 N)$ 의 연산이 필요하다. (4)식이 모든 좌표에 대해 계산되면, 임의의 영상 이미지 영역(Rectangle형)에 대한 DFT(주기는 N)는 다음과 같이, 영역의 네 꼭지점의 좌표값을 이용하여 계산할 수 있다.

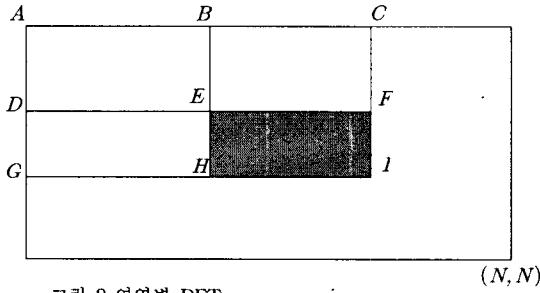


그림 2 영역별 DFT

$IFT = F(x,y,k,l)$  라고 생각할 때,

$$IFT(EFHI) = IFT(ACGI) - IFT(ACDF) - IFT(ABGH) + IFT(ABDE) \quad (5)$$

(5)의 식을 이용하여, 임의의 영역의 DFT를 구할 수 있다.

## 2.2 DFT domain에서의 패턴 매칭

2.1에서 구해진 적분형 DFT를 통해, 원하는 영역으로 DFT 영역을 제한할 수 있다. 따라서, 원하는 패턴이 영상 이미지에 존재하는지의 여부를 DFT 도메인에서 검출할 수 있다. 적분형 DFT를 이용하여 각 영역에 대한 DFT를 구한 후, 원하는 패턴의 DFT와 비교하여 그 영역에 원하는 패턴이 존재하는지를 검사한다. 입력된 영상 이미지에서의 특정 영역의 DFT를  $F_1$  이라고 하고, 원하는 패턴을  $F_2$ 라고 하면, 다음과 같이  $F_1$ 과  $F_2$ 의 닮은 패턴인지를 알아낸다.

$$\left| \frac{|F_1|}{\sum |F_1| - F_1(0,0)} - \frac{|F_2|}{\sum |F_2| - F_2(0,0)} \right| < \epsilon \quad (6)$$

(6)식이 ( $\epsilon > 0$ ) 성립한다면  $F_1$ 과  $F_2$ 는 유사한 모양을 갖는 같은 패턴이라고 생각할 수 있다.[2] 여기서,  $F_1(0,0), F_2(0,0)$ 은 각 DFT 패턴의 DC성분을 뜻하고,

$\sum |F_1|, \sum |F_2|$ 는 각 패턴의 에너지를 뜻한다.

## 3. 얼굴추출에의 적용

2D-DFT는 특정 주파수의 공간적인 위치를 알아낼 수 없다는 악점 때문에, 영상으로부터 얼굴의 위치를 찾아내는 기술인 얼굴 추출(Face Detection)에 좀처럼 이용되지 않았다. 하지만, 적분형 DFT를 이용하여 특정 주파수의 공간적인 위치를 알아낼 수 있다면, 2D-DFT 상에서 얼굴추출을 수행하는 것이 가능하다. 본 논문에서는 적분형 DFT를 이용하여, 입력된 영상 이미지에서 얼굴 패턴을 찾아내도록 한다.

### 3.1 평균 얼굴의 생성

얼굴 추출을 패턴 매칭 방식으로 수행하기 위해, 우선적으로 평균 얼굴(average face)을 구한다. 얼굴 인식(Face Recognition)의 database로 자주 이용되는 yale database 중, 정면을 바라보고 있고 얼굴 영역만 잘라낸 이미지(cropped image)만을 이용하도록 한다. 정면 얼굴만을 이용해 얻은 평균 얼굴은 다음 그림3(a)와 같다.

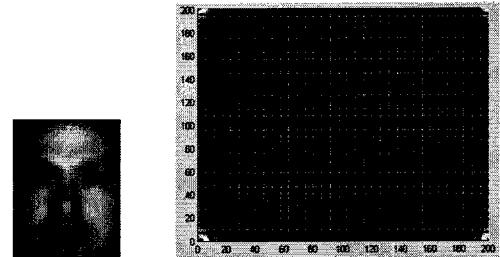


그림3(a)

그림 3(b)

그림3(b)는 그림3(a)를 DFT 변형 후의 energy를 표시한 그래프이다. 붉은 색일수록 큰 값을 갖고, 파란 색일수록 낮은 값을 갖는다. 그림에서 알 수 있듯이, 대부분의 에너지가 저주파 영역에 모여있다는 것을 알 수 있다. 이를 이용하여, (4)식을 계산 할 때, 일부 저주파의 주파수만을 취사선택하여 저장하고 연산할 수 있다. 평균 얼굴의 크기는  $28 \times 32$ 로 잡았다.

### 3.2 적분형 DFT를 이용한 얼굴추출

얼굴 추출을 위해 입력 영상의 크기를 고정하도록 한다. 입력 영상보다 작은 영상 이미지의 경우는 zero-padding을 통해 크기를 고정하도록 한다. 여기서는  $200 \times 200$  으로 입력 영상의 크기를 고정시키도록 한다. 입력 영상에서  $28 \times 32$ 보다 큰 영역을 설정, 각 영역마다 적분형 DFT를 이용하여 그 영역의 DFT를 계산하고, 평균 얼굴의 DFT와의 비교를 통해 (6)식을 만족시키는지 검사한다.  $28 \times 32$ 보다 큰 얼굴을 추출하기 위해 3번의 iteration을 들고, 각 iteration마다 평균얼굴을 1.25배 확대한다. 얼굴 검출의 과정을 요약

하면 다음과 같다.

```
get DFT of average face;
for iteration= from 1 to 4
    for x_position=0:step_size:x_size
        for y_position=0:step_size:y_size
            get DFT of this area;
            compare the DFT with face DFT pattern;
        end
    end
    enlarge face DFT pattern;
end
```

DFT를 이용한 패턴 매칭은 크기(amplitude)만을 이용할 경우 공간 변이(spatial shift)에 대해강인함을 갖는다. 따라서, 설정한 영역의 크기를 감안하여 step 크기를 정하여 스캐닝 하여도 결과는 거의 동일하다. 실험적으로 설정 영역 크기의  $\frac{1}{4}$  보다 작게 step 크기를 잡으면 얼굴 추출 성능에 큰 차이가 없었다. 일반적으로, 얼굴이 존재하는 영역 근처에서 얼굴 영역이 여러개 겹쳐서 검출된다. 유사 영역에 겹치는 부분은 (6) 식의 값이 가장 작은 영역 하나로 간주하도록 한다. 얼굴이 들어있는 영상 이미지에 적분형 DFT를 이용하여 얼굴 추출을 실험한 결과는 다음과 같다.



그림과 같이, 얼굴 영역을 추출하고 있다는 것을 알 수 있다.

#### 4. 결론

본 논문에서는 특정 영역의 주파수 성분을 계산해 낼 수 있는 적분형 DFT를 고안하였다. 적분형 DFT를 계산한 후에는 크기나 위치에 상관없이 DFT를 일정한 시간 내에 계산할 수 있었다. 적분형 DFT의 유용함을 보이기 위해 얼굴 추출(Face Detection)에 응용하여, 주파수 영역의 에너지와 패턴 매칭 방식을 이용하여 얼굴을 검출할 수 있음을 보였다. 적분형 DFT를 구하는 계산량을 더 줄일 수 있다면 더욱 유용할 것이다. 또한 주파수 영역에서 패턴 매칭을 효과적으로

할 수 있는 알고리즘을 개선할 필요가 있다. 적분형 DFT는 거의 모든 계산을 최초 입력시 수행하므로, 이 부분을 적절한 하드웨어의 힘을 빌릴 수 있다면 더욱 효과적일 것으로 생각된다.

#### 참 고 문 헌

- [1] H.M.El-Bakry,"Comments on using MLP and FFT for fast object/face detection,"in Proc. International Joint Conference on Neural Networks (IJCNN '03), vol. 2, pp. 1284-1288, Portland, Ore, USA, July 2003.
- [2] Jiangwei Li, Yunhong Wang, Tieniu Tan, A.K.Jain,"Live Face Detection Based on the Analysis of Fourier Spectra",Proceedings of SPIE, Biometric Technology for Human Identification,vol. 5404, pp.296-303, August 2004.
- [3] Alan V.Oppenheim, Ronald W.Schafer, John R.Buck "Discrete-Time Signal Processing", Second Edition, Prentice Hall.
- [4] Richard O.Duda, Peter E.Hart, David G.Stork "Pattern Classification", Second Edition, Wiley-Interscience
- [5] Katrin Meisinger, Tobias Troeger,Marcus Zeller, and Andre Kaup,"AUTOMATIC TV LOGO REMOVAL USING STATISTICAL BASED LOGO DETECTION AND FREQUENCY SELECTIVE INPAINTING", in Proc. 13th European Signal Processing Conference, September 2005.