

## Temporal 웨이브릿을 이용한 물체 추적 알고리즘

장도임\*, 김도년, 강민숙, 조동섭  
이화여자대학교 컴퓨터학과

### Object Tracking Algorithm using Temporal Wavelet

Do-Im Chang\*, Do-Nyun Kim, Min-Sook Kang, Dong-Sub Cho  
Dept. of Computer Science & Engineering, Ewha Womans Univ.

**Abstract** - Temporal 웨이브릿 변환은 2차원 이미지 시퀀스 즉, 프레임 사이의 시간적 변화를 포함한다. Temporal 웨이브릿 변환은 우선 2차원 이미지를 프레임 단위로 행방향, 열방향의 순서로 웨이브릿 변환을 수행한 후, 원영상의 1/2 해상도의 영상 프레임 시퀀스를 시간축 방향으로 웨이브릿 변환을 수행한다. Temporal Wavelet 변환의 결과로 나온 계수는 각 프레임의 특성과 프레임과 프레임 사이의 특성을 포함한다. 본 논문에서는 이 Temporal 웨이브릿 변환 결과로 나온 계수의 특성을 이용하여 물체의 움직임을 추적하는 Division Rule을 제시하고, 더욱 정확한 결과를 얻기 위해 Division Rule이 적용된 결과에 실험적으로 최적화된 가중치를 결정하여 컨벌루션을 적용한다.

## 1. 서 론

모든 형태의 영상을 분석하여 유용한 정보를 추출하는 범용 시스템의 구현은 바람직하지만 실제로는 불가능한 작업이다. 따라서 제한된 상황에서 발생할 수 있는 영상에 대한 분석을 수행하는 시스템의 구현이 적합한 접근 방법이다.

영상 분석을 위한 방법으로 이미지 변환의 결과를 이용하는 방법, 텍스처의 배치규칙을 추출하는 기하학적인 방법, 프레임 구조의 지식 표현 방법, 필터를 이용한 방법과 모델 기반 코딩을 이용하는 방법을 많이 사용한다. 이미지 변환에 이용되는 방법에는 푸리에 변환, 웨이브릿 변환 등이 있고, 프레임 구조를 이용한 방법에는 열골 영상을 분석하기 위하여 열골에 대한 모델 지식을 계층적인 프레임 지식으로 표현하는 방법[6], 동영상을 분석하여 이동 물체를 확인하고 추적하는 작업을 수행하기 위하여 문체 영역에 대한 사전 지식으로서 참조 프레임, 물체 프레임 그리고 동작 프레임으로 표현하는 방법[7] 등이 있다. 필터를 이용한 방법에는 여러 대의 카메라로부터 얻은 동영상에서 계층적 칼만 필터를 이용하여 3차원 객체를 추적하는 방법[8] 등이 있고, 모델 기반 코딩에는 Picard와 Liu가 제안한 움직이는 객체를 분할하지 않고 주기적인 움직임을 탐색하는 방법 등이 있다.[4]

본 논문에서 이용하는 방법은 웨이브릿 변환의 결과로 나온 계수를 이용하는데, 특히 Temporal 웨이브릿 변환의 결과로 나온 계수는 프레임의 특성과 프레임과 프레임 사이의 특성을 포함하므로 이 계수의 특성을 이용하여 물체의 움직임을 추적하는 알고리즘을 제시한다.

## 2. 본 론

### 2.1 웨이브릿 변환

이미지의 모양을 표현하는 스킴(Shape represent at scheme)은 다음과 같은 특징을 만족해야 한다[1].

첫째, 유일성(Uniqueness)-이미지 객체의 모양을 유

일하게 표현해야 한다.

둘째, 완전성(Completeness)-명확하고 완전한 표현어야 한다.

셋째, 불변성(Invariance under geometrical transform)-이동(translation), 회전(rotation), 스케일링(scaling)등의 기하학적인 변환에 대해서 변하지 않는 표현이어야 한다.

넷째, 민감도(Sensitivity)-비슷한 이미지들간의 차이를 쉽게 표현해야 한다.

다섯째, 추상화(Abstraction from detail)-이미지의 세부적인 사항으로부터 모양에 대한 추상적인 기본 특징까지 모두 표현할 수 있어야 한다.

웨이브릿 변환은 위에 열거된 모든 특징을 잘 만족한다. 실험결과 동일한 이미지 데이터를 이동하거나 회전, 스케일링을 해도 거의 유사한 특징 데이터를 얻을 수 있었다. 푸리에 변환(Fourier Transform)은 위의 성질을 비교적 잘 만족하지만 기본 함수(Basis Function)로 사인(sine)과 코사인(cosine)만을 사용하고 지역적인 특성이 없기 때문에 신호 데이터(Signal Data)가 불연속성을 포함하고 있거나 특별히 날카롭게 도출된 부분이 있을 경우, 이러한 지역적인 특징을 분석하기가 어렵다. 따라서 적용할 수 있는 이미지 도메인이 한정된다는 단점이 있다. 반면에 웨이브릿 변환은 기본 함수(Basis Function)로 사인(sine), 코사인(cosine) 함수뿐만 아니라 좀 더 복잡한 웨이브릿 모함수(mother wavelet function)를 사용할 수 있고, 푸리에 변환에는 없는 공간에 대한 지역적인 특성(Localized in space)을 가지고 있어서 이러한 물리적 상황에서 전체적인 특징뿐만 아니라 지역적인 특징까지도 분석이 가능하다. 또한, 웨이브릿 변환은 이미지에 대하여 다해상도(Multi-resolution) 분석이 가능하다. 즉, 웨이브릿 변환으로 나온 계수들은 이미지의 본래 해상도(original resolution)에 대하여 독립적인 정보를 제공한다.

#### 2.1.1 특징 추출

웨이브릿 변환(Wavelet Transform)은 사용하는 웨이브릿 함수의 타입(Wavelet Type)과 분해 타입(Decomposition Type)에 따라 특징 데이터 추출 성능의 차이를 보이게 된다. 또한, 변환 결과로 얻어진 수많은 계수들을 모두 특징 벡터로 사용하는 것은 매우 비효율적이므로, 절대값이 가장 큰 몇 개의 계수만을 추출하는 절단(Truncation) 작업을 하게 된다. 따라서, 본 논문에서는 다음과 같은 사항을 고려하여 특징 벡터를 추출한다.

첫째, 웨이브릿 타입 - 웨이브릿 변환은 사용하는 기본 함수(basis function)의 종류에 따라 다양한 웨이브릿 타입이 존재하게 된다. 가장 단순하고 구현이 쉬운 Harr 웨이브릿 변환과 Daubechies와 같은 다양한 종류의 웨이브릿 변환들을 사용하여 특징 데이터를 추출할 수도 있다.

둘째, 분해 타입 - 각 웨이브릿 변환에는 표준(standard) 분해와 비표준(nonstandard)분해가 존재한다. Haar 웨이브릿 변환의 경우 표준 분해 기본 함수는 사각형 모양이고, 비

본 논문은 98년도 한국학술진흥재단 학술연구비에 의한 공모과제 연구비 지원사업에 의해 지원되었음

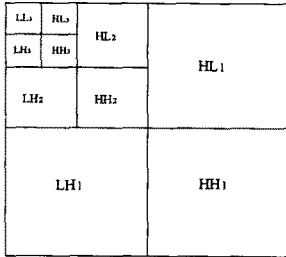
표준 분해 기본 함수는 정사각형 모양이어서 이미지가 선이나 사각형의 특징을 포함하고 있을 때는 비표준 Haar 웨이브릿 변환을 이용하는 것이 특징 추출에 효율적이다.

셋째, 절단 - 만약  $256 \times 256$ 의 해상도를 갖는 이미지 데이터에 대하여 웨이브릿 변환을 적용하면  $256^2 = 65536$ 개의 계수(coefficient)가 얻어진다. 이러한 계수들을 모두 다 특징 벡터로 사용하는 것은 계산 측면에서 상당히 비효율적이다. 따라서, 이러한 계수들 중에서 저주파 몇 개의 계수만을 취하는 절단 작업을 통하여 적절한 크기의 특징 벡터를 구성한다.

### 2.1.2 Temporal 웨이브릿 변환

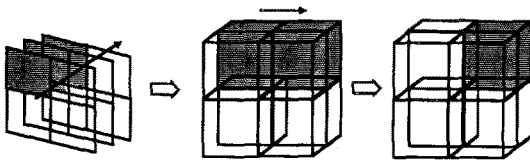
2차원 이미지에 대하여 웨이브릿 변환을 수행하면 저주파 성분이 좌측 상단에 위치한다. 저주파 성분(LL)은 원영상 이미지의 1/2 해상도를 가진 형태가 되며, HL, LH, HH 대역은 각각 수직, 수평, 대각선 방향의 고주파 성분을 가진다. [그림 1]은 2차원 이미지에 대하여 웨이브릿 변환을 세 번 실행하였을 때의 주파수 대역을 표시한 것이다.

여기에서 LL 대역을 S 컴포넌트(S Component)라 하고, 다른 대역(HL, LH, HH)을 D 컴포넌트(D Component)라 한다. 웨이브릿 변환을 반복적으로 수행하게 되면 S 컴포넌트가 다시 S 컴포넌트와 D 컴포넌트로 나뉘게 된다.



[그림 1] 2차원 이미지에 웨이브릿 변환을 세 번 적용하였을 때의 결과

이와 같은 방법으로 2차원 이미지 시퀀스에 대하여 웨이브릿 변환을 수행하면 [그림 2]와 같이 먼저 프레임 단위로 x축, y축 웨이브릿 변환을 수행한 후 그 결과를 시간축(t축)으로 다시 웨이브릿 변환을 수행한다.



[그림 2] Temporal 웨이브릿 변환

### 2.2 물체 추적 알고리즘

제안하는 물체 추적 알고리즘은 다음과 같다.

**Algorithm 1** : 물체추적 알고리즘(Object Track Algorithm : OT 알고리즘)

Step 1 : 이미지 시퀀스에 대하여 Temporal 웨이브릿 변환을 하고 그 결과를 이미지 시퀀스로 저장하여 S 컴포넌트(S Component) 시퀀스를 구한다.

Step 2 : Division Rule에 따라 4개의 S 컴포넌트 시퀀스( $S_0$ )를 선택한다.

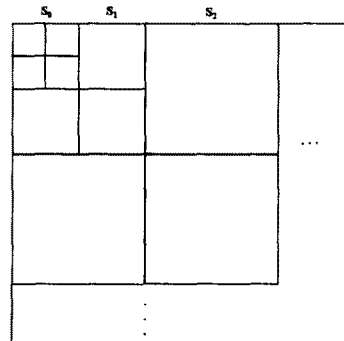
Step 3 : 선택된 4개의 S 컴포넌트에 대해 실험적으로 최적화된 가중치를 결정하여 컨벌루션

(convolution)을 하여 하나의 컴포넌트로 만든다.

Step 4 : 컨벌루션 결과를 웨이브릿 역변환하여 결과 이미지를 얻는다.

### 2.2.1 Division Rule

Temporal 웨이브릿의 결과로 만들어진 이미지 시퀀스를 컨벌루션하여 움직임을 추적할 때, 특성이 많이 포함된 부분을 선택하는 것이 중요하다. 따라서 저주파 성분을 포함하는 S 컴포넌트를 선택하게 되며, 선택한 S 컴포넌트로 움직임 추적이 어려운 경우 그 개수를 늘려 나가는데 다음과 같은 규칙을 따른다.



[그림 3] Division Rule

### Rule 1 : Division Rule

$$S_0 = 4$$

$$S_K = 2^{K+1}, K=1,2,3,\dots$$

$$\sum_{K=0}^n S_K = N = 2^{n+2}$$

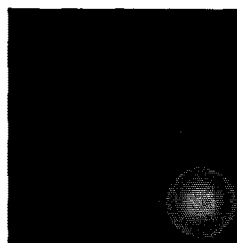
### 2.2.2 Convolution을 위한 가중치 선택

S 컴포넌트들 사이에서 특징을 포함하는 정도가 다르기 때문에 선택된 컴포넌트를 컨벌루션할 때 가중치를 두어 더욱 정확하고 뚜렷한 결과를 얻을 수 있다.

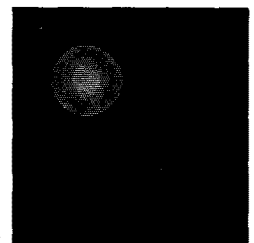
$S_0$ 이 S 컴포넌트 A, B, C, D로 구성되어 있다면, 각각의 컴포넌트는 2차원으로 구성된 (i, j)의 웨이브릿 계수를 갖는다. 각각의 컴포넌트 원소를  $a_{ij}$ ,  $b_{ij}$ ,  $c_{ij}$ ,  $d_{ij}$ 라고 표시한다. 이와 같이 구성된  $S_0$ 에 가중치를 0.2, 0.4, 0.1, 0.3으로 결정하면 컨벌루션의 결과 컴포넌트 R은  $0.2a_{ij} + 0.4b_{ij} + 0.1c_{ij} + 0.3d_{ij}$ 가 된다.

### 2.3 실험 및 결과

실험에 사용한 16개의 프레임으로 구성된 이미지 시퀀스에는 구(球)가 직선운동을 하는 움직임이 포함되어 있다. 프레임1에서 프레임16으로 움직이는 이미지 시퀀스를 입력으로 가진다. 다음은 16개의 이미지 중 처음과 마지막 프레임이다.

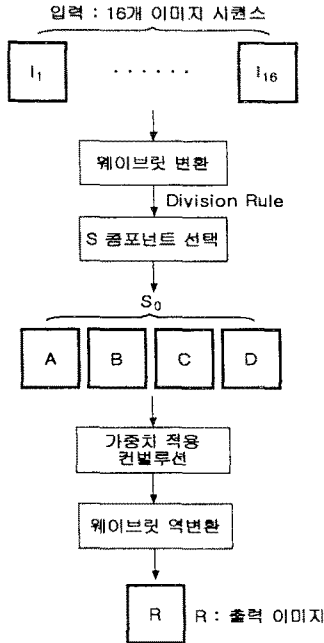


[그림 4] 프레임 1



[그림 5] 프레임 16

[그림 6]은 물체 추적 알고리즘 실험의 전체적인 block diagram이다.



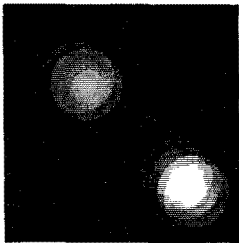
[그림 6] 제안한 물체 추적 알고리즘의 전체적인 진행도

### 2.3.1 Temporal 웨이브릿 변환

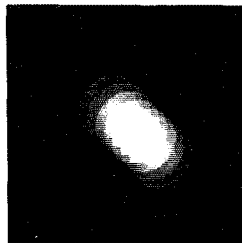
16개의 이미지를 Temporal 웨이브릿 변환하여 4개의 S 성분과 12개의 D 성분으로 구분한다.

### 2.3.2 S0(4개의 S Component) 선택

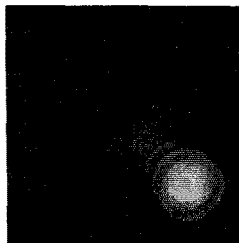
다음 그림은 Temporal 웨이브릿의 결과에서 선택한 4개의 S 성분이다.



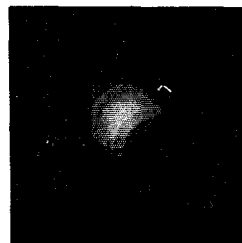
[그림 7] S 성분 -A



[그림 8] S 성분 -B



[그림 9] S 성분 -C

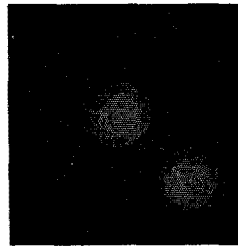


[그림 10] S 성분 -D

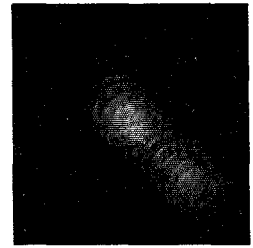
### 2.3.3 컨벌루션 (Convolution)

S 성분 A, B, C, D에 가중치를 적용하여 컨벌루

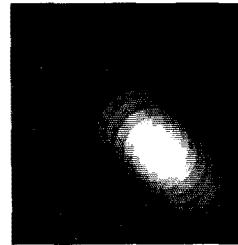
션하여 웨이브릿 역변환한 결과는 다음과 같다.



[그림 11] 가중치 0.25, 0.25, 0.25, 0.25를 적용한 경우



[그림 12] 가중치 0.4, 0.3, 0.2, 0.1을 적용한 경우



[그림 13] 가중치 0.875, 0.125, 0.125, -0.125를 적용한 경우

## 3. 결 론

Temporal 웨이브릿 변환의 결과로 나온 계수의 특성을 이용하여 물체의 움직임을 추적하는 Division Rule을 제시하였고, Division Rule이 적용된 결과에 실험적으로 최적화된 가중치를 결정하여 컨벌루션을 적용하였다.

실험결과 Division Rule이 적용된 결과에 대해 컨벌루션을 하였기 때문에 적은 프레임을 가지고 연산을 할 수 있었고, 가중치를 주어 컨벌루션 계산을 하여 움직임을 정확하게 추적할 수 있었고, 뚜렷한 결과를 확인할 수 있었다.

### [참 고 문 헌]

- [1] Joannis Pitas. "Digital Image Processing Algorithm." Prentice Hall, 1993.
- [2] Hwnag V. S., L. S. Davis and T. Matshyama, "Hypothesis Integration in Image Understanding system," Computer Vision Graphics and Image Processing, Vol.36, pp.321-371, 1986.
- [3] C. Wang and S. Srihari, "A Framework for Object Recognition in a Visually Complex Environment and its Application to Locating Address Brocks on Mail Pieces," International Journal of Computer Vision, Vol.2, pp.125-151, 1988.
- [4] Fang Liu, "Modeling Spatial and Temporal Textures", MIT Press, Ph.D Thesis, 1997.
- [5] H. V. Jagadish, "A Retrieval Techniques for Similar Shapes," Proceedings of the ACM SIGMOD International Conference on the Management of Data, pp.208-217, 1991.
- [6] 김봉근, 최형일, "얼굴 영상 분석 시스템," 한국 인지과학회 논문지, Vol.3, No.1, pp.79-111, 1991.
- [7] 이진호, 최형일, "동작 가설 생성과 검증을 통한 이동 물체의 검출 및 추적," 전자공학회 논문지, Vol.30-B, No.8, pp.41-53, 1993.
- [8] 정순기, 원광연, "계층적 칼만 필터를 이용한 관절을 가진 물체의 추적 및 운동 추적," 정보과학회 논문지, 제23권 11호, pp.1105-1117, 1996.