

《主 題》

초저속 전송을 위한 wavelet변환기반의 동화상 압축기술

김 성 환 · 이 홍 규

(한국과학기술원 전산학과)

■ 차 례 ■

I. 서 론

II. 초저속 영상통신 표준화 동향

III. Wavelet 변환

IV. Packet Wavelet 변환

V. Wavelet 변환을 이용한 동화상 압축기법

VI. 제안된 동화상 압축기법

VII. 결 론

ABSTRACT

This paper presents a survey of video coding schemes which use wavelet transform for the videophone on very low bit rate communication channel(eg. 10 Kbps Public Service Telephone Network). Firstly, we introduce the standardization efforts to make the low bit rate videophone architecture and the typical application of low bit rate video coding scheme. Secondly, we summarize the several requirements on videophone, delay, encoder/decoder complexity, low bitrate, and progressive transmission capability. Third, we review the basic theory of wavelet transform without much mathematics. We compare the wavelet transform with short-time fourier transform and subband filters. Fourth, we summarize the video coding schemes proposed so far, and evaluate them with the requirements. Lastly, we conclude with future research directions.

keyword : wavelet transform, videophone, very low bit rate

I. 서 론

일상생활에 있어 대부분의 정보는 시각에 의해 얻어지고 있으므로 영상 정보를 교환할 수 있는 영상 통신에 관한 연구가 오래전부터 연구되어져왔다. 진화망을 이용한 비디오폰의 경우, 한 프레임이 가로 176, 세로 144 pels로 구성되고 초당 10 frame을 전송한다고 할 때, 약 2 Mbps의 정보량을 갖게 되며, 전화망의 채널폭은 약 10 Kbps이므로 200 대 1 정도의 높은 압축율을 요하게 된다. 따라서 영상부호화 기법의

개발은 필수적이라 할 수 있다. 최근 들어 컴퓨터, VLSI 등 관련분야의 발전과 더불어 ISDN, teleconferencing 및 videophone, HDTV등 영상 통신에 대한 수요가 커짐에 따라 영상 정보를 이용하는 방법들에 대한 연구가 촉진되고 있다.

본 기고에서는 세계적으로 무척 활발하게 연구되고 있는 초저속 영상통신을 위한 동화상 압축기법의 현재의 연구현황을 다루기로 한다. 여기서 초저속이라는 말은 N-ISDN(narrow ISDN)이나 B-ISDN(broad ISDN)망과 같은 고속통신망과 대별되는 기존의 전

화망(public service telephone network)이나 무선망을 지칭한다. ISDN망이 보편화된다고 할 때 이러한 초저속망을 대상으로 한 압축기법의 유용성에 대해 의문을 제기할 수도 있지만, 비용상의 문제로 또한 무선망과 같은 경우 근본적인 대역폭의 제한과 다양한 사용자의 멀티미디어에 대한 요구의 증가는 초저속 영상압축 기법을 연구하는 필연적인 이유가 된다.

본 기고에서는 다음과 같은 순서로 초저속 영상통신의 기술을 소개하고자 한다. 우선 2장에서는 ITU에서 제시된 자료를 바탕으로 초저속영상전송을 위한 대표적인 응용분야와 함께 요구사항을 분석하고 현재 ITU를 중심으로 한 초저속 영상압축 기법의 연구현황을 소개하고, 3장에서는 wavelet변환의 기본적인 사항을 다루며, 기존의 신호처리방법인 STFT(short time fourier transform) 및 SBC(subband coding)과 비교결과를 제시한다. 4장에서는 wavelet변환을 중심으로 한 지금까지 제시된 여러 비디오코딩기법들을 제시된 요구사항 측면에서 비교한다. 마지막으로 5장에서는 초저속영상 압축안의 방향에 대한 견해와 함께 결론을 맺는다.

II. 초저속 영상통신 표준화 동향

현재 초저속 영상통신을 위한 표준화 협의는 ITU-T SG15/LBC회의를 중심으로 이루어지고 있으며 ISO 산하의 MPEG-4그룹과도 밀접한 협력 아래 표준안을 진행중이다. 이 장에서는 초저속 영상통신의 주 응용분야를 분류하고 각 응용분야가 요구하는 공통적인 요구사항들을 제시하기로 한다. 또한 얼마전 열린 LBC회의의 토의사항을 소개하고, 현재 ITU를 중심으로 연구되고 있는 다양한 영상압축방법에 대해 간단하게 소개하고자 한다.

다음 절에서는 초저속 영상통신기술의 응용분야와 각 응용분야에서 총괄적으로 요구되는 요구사항들을 정리하기로 한다.

2.1 초저속영상통신의 응용분야와 요구사항

ITU에서는 초저속 영상통신의 응용분야를 다음 표 1와 같이 세 분류로 나누고 있다. 대체적으로 대화형, 멀티미디어통신을 지원하며, 전송속이 작기 때문에 에러수정부호화 등의 부단이 크므로 되도록이면 부호화 기법 자체가 전송요류에 강하게 설계되어야 한다.

다음은 위의 응용분야에서의 요구사항을 ITU의 자료를 바탕으로 정리한 것이다. 결국 초저속영상통신을 위한 영상압축기법의 설계는 서로 상충되는 지연시간, 복잡도, 전송률 측면을 조절하는 것이며, 여기에 점진적전송 능력을 참가함으로써 packet망에서의 전송요류나 검색(browsing)응용에 효과적으로 적용할 수 있다.

1. 지연시간(delay) : 최대 지연시간은 응용에 따라 다르지만 일반적으로 대화형의 경우는 encoder의 경우, 400ms이하이며, decoder의 경우 50ms에서 1000ms이며, 초기지연은 최소가 되도록 한다.
2. 복잡도(complexity) : encoder의 경우는 비교적 낮게, decoder의 경우는 최대한 낮게 잡는다.
3. 전송률(bit rate) : MPEG-4의 전송율은 64Kbps 이하로 하고 있으나 무선망등을 고려해볼때 10Kbps이하에서도 동작하여야 한다.
4. 점진적 전송(progressive transmission) : 다양한 복잡도의 decoder및 응용을 고려해볼때 낮은 해상도와 높은 해상도를 동시에 전송할 수 있어야 한다.

표 1. 초저속 영상통신의 주응용분야

응용분야	응용 분야의 특성	전송 매체
General	Real-Time Mainly symmetric Error Resilience가 중요	(1B) ISDN PSTN Mobile Net(GSM, DECT, UMTS, FLMPST)
Videophone	Head and Shoulder image Mainly symmetric Error Resilience가 중요	(1B) ISDN PSTN Mobile Net(GSM, DECT, UMTS, FLMPST)
Others	videoconferencing, tele-working, multipoint videotelephony, remote expert, remote classroom electronic news gathering	(1B) ISDN PSTN Mobile Net

이러한 요구사항들을 배경으로 최근 초저속 영상 통신에 대한 수요증가로 단시간에 표준을 만들 필요가 있으며, 이러한 추세에 따라 ITU에서는 LBC회의를 중심으로 단기 표준안을 만들고 있으며, 보다 나은 화질을 위한 장기표준안을 병행하고 있다. 다음 절에서는 현재 LBC회의를 통해 진행되고 있는 표준화동향을 살펴보기로 한다.

2.2 최근 표준화 회의의 결과

다음은 1994년 3월 28일에 프랑스 텔레콤 주관으로 파리에서 열린 ITU T SG15/LBC회의에서 결성된 개략적인 사항이다. 이번 회의에는 20여개 국가(약 40개 업체)에서 100여명이 참석하였으며, 주로 기간통신업자들이 중심이 되었다. LBC회의의 각 부문별 의장과 부문별 제안서류는 다음 표와 같다.

부 분	의 장	권 고 안
시스템 및 요구사항	R.Schaphorst, (Delta Info Sys)	H.32P
영상부호화	K. Rijkse, (Netheland Telecom)	H.26P, H.26P/L
음성부호화	R. Cox, (AT&T)	AV.25Y
다중화 및 시스템제어	R. Brainard, (AT&T)	H.22P, H.24P H.DLP

각 subgroup에 대한 개괄적인 사항을 표2에 정리해 놓았다. 본 기고에서는 초저속영상통신을 위한 영상 부호화기법에 한하여 논의를 한정하기로 한다. 다음은 현재 초저속영상통신을 위한 단기 및 장기 표준으로 연구되고 있는 주제들을 정리해보기로 한다.

2.3 표준으로 고려되는 영상부호화기법들

현재 초저속 영상통신용으로 제시된 영상 format은 H.261과의 호환성을 고려하여 QCIF로 통일하는 것으로 결정이 되었지만 응용분야에 따라서 QCIF보다 작은 해상도를 가진 QCTX를 쓰기도 한다. 다음은 QCIF와 QCTX에 대한 간단한 설명이다.

	QCIF	QCTX
No. of luminance line	144	112
No. of luma pixel/line	176	128
No. of chroma lines	72	56
No. of chroma pixels/line	88	64
Total no of pixels/frame	38016	21504
No. of macroblocks/frame	99	56

현재 ITU에서는 초저속 영상통신을 위해 단기표준과 중/장기표준으로 나누어 표준안을 진행하고 있다. 표3은 현재 진행되고 있는 영상부호화 기법들을 영상

표 2. ITU 표준회의의 중요한 결정사항들

부 분	토의된 내용
시스템 및 요구사항	ISDN용 비데오폰 표준인 H.320과 PSTN 표준인인 H.320P와의 interoperability에 대한 많은 연구가 행해졌다. CNET에서는 PC상에서 구현 가능한 TMN-1을 기반으로 한 multimedia terminal을 하드웨어로 구현하여 태모를 보여주었다. 영상과 음성신호를 multiplexing할 수 있는 optional data channel을 정의하기 위해, LBC expert group에서 recommendation H.DPL을 개발하도록 하였다. 이 data channel을 이용하면, 기존의 비디오폰을 이용하여, remote database의 정보를 access할 수 있게 한다. DTE/DCE partitioning within H.32P Draft Recommendation H.32P가 작성되었다.
영상부호화	H.26P부호화 알고리즘에 대한 reference test model(TMNI : Near Term Test Model)이 정의되고 이것은 여러가지 실험에 표준모델로 사용되었다. H.261 QCIF가 mandatory 표준 영상 format으로 결정되었다. 그러나 각각의 측면을 고려하여 QCIF bitstream의 일부만을 decoding하는 방안이 검토되고 있다. 16x16, 8x8, 16x8, 8x16 motion estimation의 성능 차이가 크지 않으므로 기존의 H.261과의 호환성을 고려하여, 16x16 motion estimation을 수용하기로 합의하였다. chrominance motion estimation의 경우에는 half-pel 단위의 motion estimation을 사용하기로 하였다. Draft Recommendation H.26P가 작성되었다.
음성부호화	다섯개의 후보 음성부호화기의 성능이 몇 가지 주관적인 테스트를 통하여 평가되었다. 이 중 두개의 부호화기(프랑스 텔레콤, DSP Group)의 부호화기만이 다음 테스트를 위한 부호화기로 선택되었다. 이들은 각각 6.8Kbps와 6.4Kbps에서 동작한다. 이들 두 부호화기의 성능은 G.728 부호화기의 성능에 매우 근접하다. Draft Recommendation AV.25Y가 작성되었다.

다중화	전송형식은 frame synchronous구조 대신 packet 형식으로 하기로 합의하였다. 이 형식은 여러가지 multimedia terminal에서 장점을 갖는다. Multiplexing요구사항으로 efficiency, delay, flexible/extendible, error control에 대한 것으로 최종적으로 packet length, packet overhead, error management, flexible syntax의 feature로 변환되었다. LAP-F와 같은 표준 LAP 프로토콜이 사용된다. Draft Recommendation H.22P가 작성되었다.
시스템 control	Supervisory channel이 여러가지 다양한 목적으로 사용된다. Draft Recommendation H.24P가 작성되었다.

표 3. 영상모델에 따른 영상압축안

level	영상모델	부호화되는 정보	부호화 방법	비고
1	pels	color of pels	PCM	
2	statistically dependent pels	color of pels or block of pels	predictive coding transform coding	
3	translationally moving blocks	color of blocks and motion vectors	motion compensated hybrid DPCM/DCT coding	H.261 변형 TMN-2
4	moving structures	moving parameters or shape and motion	fractal coding contour texture coding	
5	moving unknown objects	shape, motion and color of each object	analysis/synthesis coding	
6	moving known objects	shape, motion and color of the known objects	knowledge based coding	
7	facial expressions	action units	semantic coding	

모델을 기준으로 나눈 것이다.

level은 영상에 대한 가정(assumption) 정도를 나타내며, level이 높아질 수록 영상에 대한 가정이 많아지게 되고 어떤 의미에서는 인간의 지각특성을 고려한 압축안이라고 할 수 있다. 처음 세개의 모델은 현재 MPEG 계열의 표준안에서 쓰이고 있는 방법이며, MPEG-4의 단기표준안으로 고려하고 있는 영상모델로써 영상을 구성하는 각 화소를 단순히 디지털화된 신호 또는 신호간에 1차 또는 m차 상관 관계를 가지고 있다는 식으로 모델링된다. Level 4와 5의 방법은 프랙탈과 컴퓨터비전기술이 영상부호화에 적용된 것이며, fractal coding, segmentation based coding(region based coding, object oriented coding)등으로 분류되어 연구되고 있으며, level 6과 7의 방법은 인간의 표정을 형성하는 안면근육의 기본동작을 추출하고 영상에서 표정의 변화를 감지하여 전송하는 방법으로 알고리즘이 실시간에 구현이 어렵고 영상에 대해 너무 많은 가정을 했다는 단점으로 현재 MPEG-4의 장기계획안으로 설정되어 있으며, model based coding이라는 분야로 연구되고 있다. Level 4 이후는 영상을 단순한 화소의 모임으로 보는 level 1-level 3과 구별하여 2세

대 영상압축기법이라는 용어를 사용하기도 한다. 참고로 wavelet변환을 중심으로 하는 영상압축안은 중장기 표준안의 한 형태로서 제시될 전망이다.

다음 절부터는 각각의 안에 대하여 간단히 살펴보기로 한다.

2.4 TMN-2(Near-Term Test Model-2)

10 Kbps급 동화상 부호화 시스템의 test model인 TMN-2는 DCT 기반의 H.261를 초저속망에 맞도록 수정한 형태를 띠고 있으며, 기본적인 차이점은 다음과 같다. 첫째, 밝기성분에 대해서는 정수 단위의 움직임 추정을 수행되고, 색성분에 대해서는 반화소 단위의 움직임 추정을 수행한다. 또한, 루프필터(loop filter)는 사용되지 않는다. 둘째, 움직임 추정은 블록 또는 매크로 블록 단위로 수행한다. 셋째, GOB(group of block)계층은 존재하지 않으며, 양자화 스텝 사이즈는 매크로 블록 단위로 전송된다. 또한 각 매크로 블록의 전송여부를 나타내는 1비트 시그널을 매크로 블록 계층마다 전송한다.

DCT가 가지는 근본적인 화질저하요인으로 초저속 영상통신에 적용하기엔 문제점이 많지만 MPEG과

H.261의 흐름이 DCT를 중심으로 개발되어 왔고 다양한 하드웨어가 존재하기 때문에 단기표준안으로 연구되고 있다. 지연시간, 복잡도에서는 양호하지만 비트율과 점진적 전송이라는 측면에서는 성능이 떨어진다.

2.5 프랙탈을 이용한 동화상 압축기법

프랙탈 영상압축에서의 문제는 어떤 영상에 대하여 가장 적은 량의 비트를 가지고 가장 근접하게 영상을 기술하는 IFS(iterated function system)의 계수를 추정하는 문제이다. 이 문제는 다음 두가지를 가지는 썬이다. 첫째는 대부분의 자연영상은 부분적인 구조를 전체적으로 유지하는 형태이며(예를 들어 구름의 경우) 둘째, 주어진 영상을 몇 개의 변환과정만으로 영상에 근접하게 attractor를 만들어낼 수 있는 IFS를 찾을 수 있다는 두 가지 조건이며, 대부분의 자연영상에 대해서 합당한 조건이 될 수 있다[Jacq 92].

프랙탈 영상압축은 다음의 기본적인 정리인 축소변환(contractive mapping) 정리와 콜라주 정리로 기판으로 한다.

[축소변환 정리]

임의의 두 영상 X와 Y의 차이를 $d(X, Y)$ 라고 하고 축소변환 W는 다음을 만족해야 한다. 0이상 1미만의 실수 s에 대하여 $d(W(X), W(Y)) \leq sd(X, Y)$. 위의 조건을 만족하면 W는 다음과 같은 조건을 만족하게 된다.

$$W(X_n) = X_n$$

$$\lim_{n \rightarrow \infty} W^n(Y) = X_n$$

[콜라주 정리]

$$d(X, X_n) \leq \frac{1}{1-s} d(X, W(X))$$

여기서 X_n 는 W가 정해지면 W에 종속되는 IFS의 최종영상을 의미하며 다른 말로는 attractor 영상이라고 한다. X는 물론 원영상이며, $W(X)$ 는 X영상에 W라는 축소변환을 한번 적용시킨 결과이다. 다시 말하면 원영상과 X_n 즉 수신측에서 복원된 영상의 차이는 오른쪽에 제시된 식에 의해 한정(bound)될 수 있다는 의미이며 결국 오른쪽의 식을 최소화하는 W를 찾는 것이 프랙탈 압축의 핵심이 된다.

여기서 W를 찾는 과정은 영상을 임의 크기의 블록으로 나누고 주어진 영역 안에서 affine similarity와 contrast/luminance를 만족하는 블록을 찾는 과정으로도 가능하지만 다른 방법도 가능하다.

프랙탈 기법을 동화상 압축에 적용하는 방법은 여러가지가 있을 수 있지만 대표적인 방법은 단순히 3차원으로 확장하여 3차원 블록간의 유사성을 찾을 수 있으며, 다른 방법은 기존의 움직임 추정기법과 결합하여 적용하는 방법 또는 texture부분에 대해서만 적용하는 방법이 제안되고 있다. 아직 동화상압축에서는 좋은 성능을 보이지 못하고 있으며, 지연시간, 복잡도, 점진적 전송측면에서 좋지 못하지만 비트율 측면에서 향상이 가능하다.

2.6 모델 분석기법을 이용한 동화상압축기법

개인상호간의 대화에 있어서 언어에 의한 정보전달이 40%정도이며, 얼굴표정 및 머리의 움직임, 제스처 등을 비롯한 비언어에 의한 정보전달이 60%에 달한다는 사실은 이제 통신이 음성의 정확한 전달 뿐만 아니라 비언어 역시 잘 전송할 수 있어야함을 나타내준다[kim 93].

기존의 영상부호화 방식은 동화상이 통계적인 성질을 갖는 2차원 연속신호로 모델링하여 화소간의 상관 및 진폭분포, 주파수의 에너지분포 등 통계적인 지식을 이용하여 영상을 압축하고 있다. 그러나 초저속용이 되기 위해서는 높은 압축비가 필요하며, 영상에 대한 가정이 없이는 만족할 만한 성능을 얻을 수 없다. 모델 분석 동화상 압축방법에서는 초저속용 동화상이 얼굴과 어깨만으로 구성된 영상이라고 가정하며 얼굴동화상을 2차원 통계적인 신호라는 입장에서 탈피하여 3차원 얼굴모델의 2차원 화면 상의 투영이라고 보고 얼굴과 어깨를 표현한다. 다시 말하면, 얼굴과 어깨로 구성된 동화상에서 미리 가지고 있는 얼굴의 3차원 모델을 정합하여 얼굴의 변화를 파라미터화하고 구해진 파라미터를 송신한다. 이때 파라미터는 얼굴의 변화, 머리의 동작을 나타낸다. 수신측에서는 파라미터를 수신하여 미리 가지고 있는 얼굴의 3차원 모델에 파라미터를 가하여 모델을 변형시킨 다음 주 머리를 이동한다던가 얼굴 표정을 옮긴 모습에서 옷은 형태로 바꾼 후에 컴퓨터 그래픽을 이용하여 얼굴과 옷색 등을 칠한다.

정리하면 송신측에서는 입력영상에 대하여 다음과 같은 일을 하게 된다. 첫째, 얼굴 및 상반신의 3차원 움직임 추정을 하게 된다. 둘째, 표정분석이 이루어지

표 4. 모델기반부호화와 기존의 방법과의 차이점

	종래의 영상부호화	모델분석 부호화
부호화에 이용하는 지식	진폭분포, 화소간 상관 등의 통계적 지식	부호화대상의 3차원 모델 등의 영상정보
전송정보의 내용	과형을 표현하는 신호값의 표현	부호화대상의 3차원 운동 변형 등의 분석, 인식결과
충실도의 기준	과형 값의 자승오차	두부의 운동, 표정변화등의 분석, 인식대상요소의 재현성

는데 미리 가지고 있는 3차원 얼굴모델을 변형시켜 봄으로써 구할 수 있다. 여기서 재미있는 사실은 인간의 표정이 무척 다양하고 많아보이지만 사실은 64개의 기본적인 안면근육의 움직임의 합으로 표현할 수 있다는 사실이다. 셋째, 기준영상에 대해 위의 분석이 이루어지면 다음에 오는 영상에 대해서는 얼굴 및 상반신의 3차원 모델을 갱신하는 작업으로 내치될 수 있다.

반대로 수신측에서는 다음과 같은 일을 한다. 송신측으로부터 얼굴의 운동정보, 상반신의 운동정보, 표정정보, 갱신정보를 받은 후에, 배경을 생성하고, 얼굴의 3차원 움직임을 합성하고 표정정보로부터 얼굴 표정을 합성한 후에 모델 갱신정보를 통해 이번 영상에 대한 3차원 얼굴모델을 완성한 다음 2차원에 투영한다.

마지막으로 종래의 영상부호화 즉 변환기반의 영상부호화와 모델분석부호화를 비교한 결과는 다음과 표4와 같다.

모델기반 압축안은 지연시간, 복잡도, 점진적 전송 측면에서 약점을 가지며, 비트율 측면에서는 상당한 감소가 이루어지지만 대상영상에 대한 너무 많은 가정이 문제가 될 수 있다.

다음 장에서는 중장기안의 하나로 제시되고 있는 wavelet변환을 이용하는 동화상 압축기법에 대한 설명을 한다.

III. Wavelet 변환

Wavelet 이론은 응용수학에서 처음 소개된 후 최근 컴퓨터 비전 분야에서 연구되어 온 다중해상도 표현과 연관성이 있음이 밝혀졌으며 이산 wavelet 이론은 이산신호의 subband 분해방법과도 연관성이 존재한다. 이 절에서는 wavelet이론에 대해 설명하고, 이산 wavelet변환과 다중해상도 표현 및 서브밴드분해방법과의 연관성을 설명한다.

3.1 wavelet변환

Wavelet이란 $\psi(x)$ 로 정의되는 mother wavelet을 변이시키고 확대/축소시킴으로써 얻어지는 함수들의 집합을 말하며, 어떤 신호 $f(x)$ 의 wavelet변환은 다음과 같다.

$$Wavef(a, b) = \frac{1}{|a|^{1/2}} \int \psi\left[\frac{x-b}{a}\right] f(x) dx \quad (1)$$

여기서 ψ 는 mother wavelet(a prototype wavelet)이며 매개 변수 a, b 가 실수($a, b \in R$ with $a \neq 0$)일 때 연속 wavelet 변환이라 하고, a, b 가 정수($a = a_0^m, b = n b_0 a_0^m$, with $m, n \in Z$ and $a_0 > 1, b_0 > 0$ fixed)일 때 이산 wavelet 변환이라고 한다[Morl 82]. 특히 a_0 가 2이고 b_0 는 1일때 직교정규 기저를 만들 수 있다.[Daub88a]

매개변수 b 는 시간축 상에서의 변의를 나타내며, b 를 변화시킴으로써 wavelet 기저를 원하는 곳에 놓을 수 있다. 매개변수 a 는 크기인사이며 wavelet기저의 크기를 조정한다. 그림 1(b)에서 보듯이 a 가 작으면 시간축 상에서 좁은 구역에 놓이게 되며 Heisenberg의 불확정성 원리에 의해 주파수 축에서는 넓은 영역을 차지한다. 따라서 ψ 를 고 주파수 대역에서는 세밀한 해상도를 가지게 하고 낮은 주파수에 대해서는 상대적으로 나쁜 시간 해상도를 가지게 할 경우 긴 저주파수 성분에서의 감사가 나타나는 고 주파수성분이 섞인 신호의 경우에도 쉽게 처리할 수 있다. 따라서 자연영상의 경우 화소의 값이 천천히 변해가며, 어느 순간 물체의 가상자리 부분에서 화소의 값이 바뀔을 생각하면 wavelet변환을 영상 압축에 이용함으로써 윤곽을 또렷이 살리면서도 최대한 압축이 가능함을 알 수 있다. 반면에 windowed fourier 변환(그림1(a)의 경우 시간-주파수 공간의 해상도가 고정되게 되고 따라서 nonstationary한 신호의 경우 처리가 어렵다. 영상 압축에 쓰이고 있는 DCT(discrete cosine transform)의 경우 windowed fourier 변환으로 볼 수 있으며, 신호의 특성이 불분명한 예측 오차 영상 압축에는 한계점이 드러난다.

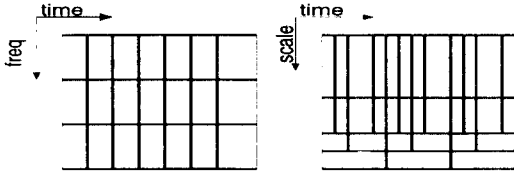


그림 1. Time-Frequency 영역에서의 Windowed Fourier 변환과 Wavelet 변환의 해석 [Daube 90]

3.2 Orthogonal Wavelet 변환

컴퓨터 비전에서는 여러가지 해상도로 영상을 분석하는 것이 필요하며, 효율적인 처리를 위해 wavelet 변환이 적용될 수 있다 [Mall 89a]. 영상압축 분야에서도 마찬가지로 최종사용자가 사람이라는 점에서 인간 시각시스템의 여러 특성을 이용해야 하며 그 중 중요한 특성이 다중해상도 또는 피라미드 식의 영상 처리이다.

다중해상도 분석은 다음과 같이 수학적으로 정의될 수 있다.

[정의 2.3] 어떤 함수 f 의 다중 해상도 분석은 함수 f 를 서로 다른 scale에서의 근사 함수(approximation) 집합의 계층구조(hierarchy)로 구성하는 것이다. 즉, square integrable 함수 공간 $L^2(R)$ 공간을 다음과 같은 부분 공간 V_j 의 연쇄로 나누고 아래의 계층구조중 V_m 에 직교정규 투영하면, 해상도 2^m 에서의 함수 f 의 근사 함수가 된다.

$$\dots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots \tag{2}$$

이때 $\bigcap_{j \in Z} V_j = 0, \bigcup_{j \in Z} V_j = L^2(R)$

여기서 $V_m \rightarrow L^2(R)$ for $m \rightarrow \infty$.

다중 해상도 분석에서의 scaling은 다음 조건에 의해 표시된다.

$$g \in V_j \Leftrightarrow g(2^{-j} \cdot) \in V_{j+1} \tag{3}$$

다중 해상도에서의 translation invariance는 다음 조건에 의해 표시된다.

$$g \in V_0 \Leftrightarrow g(\cdot - k) \in V_0 \quad \forall k \in Z \tag{4}$$

다중 해상도 분석이 되기 위한 마지막 조건으로 V_0 에 속한 ϕ 의 integer translate $\phi(\cdot - k)$ 가 V_0 의 직교정규 기저를 형성해야 한다. 즉,

$$\forall g \in V_0, \int |g(x)|^2 dx = \sum_k \left| \int g(x) \phi(x - k) dx \right|^2 \tag{5}$$

[정의 2.4] 어떤 함수 f 의 단계적인 근사 함수는 부분 공간 V_j 에 그 함수를 직교 투영한 것으로 정의한다.

$$P_j f = \sum_{k \in Z} \phi_{jk} \langle \phi_{jk}, f \rangle$$

또한 일정한 함수를 단계적 근사함수들의 계층구조로 다중해상도를 표현할 수도 있지만 이웃하는 두 계층에서의 근사 함수 $P_j f$ 와 $P_{j+1} f$ 의 정보 차이 V_{j+1} 를 유지함으로써도 표현이 가능하다. 즉 공간 V_j 와 직교 complement인 공간 W_j 에 f 를 직교 투영한 $Q_j f$ 의 함으로 표현할 수 있다 [Daub 90]. 즉,

$$W_j \perp V_j,$$

$$V_j \oplus W_j = V_{j+1},$$

$$Q_j f = P_{j+1} f - P_j f.$$

$Q_j f$ 를 직접 구하기 위해서, 먼저 W_j 공간의 직교정규 기저를 찾아야 한다. 여기서, 4개의 다중 해상도 분석 요구 조건(2.2)(2.3)(2.4)(2.5)으로부터 다음의 W_j 에 대한 식들이 유도된다.

$$g \in W_j \Leftrightarrow g(2^{-j} \cdot) \in W_0 \tag{6}$$

$$g \in W_0 \Leftrightarrow g(\cdot - k) \in W_0 \tag{7}$$

$$W_j \perp W_{j'} \quad \text{for } j \neq j' \tag{8}$$

$$\bigoplus_{j \in Z} W_j = L^2(R) \tag{9}$$

$$\forall g \in W_0, \int |g(x)|^2 dx = \sum_{k \in Z} \left| \int \psi(x - k) g(x) dx \right|^2 \tag{10}$$

식(2.6)은 W_j 가 W_0 의 scaled version이라는 점을 나타내주며, W_j 가 이산(discrete) translation $2^{-j} Z$ 에 대

해 translation invariant하다는 점은 식(2.7)에서 알 수 있다. (2.8)과 (2.9)는 W_j 공간들이 서로 직교하고, W_j 공간의 합이 $L^2(R)$ 임을 나타낸다. 이때 (2.10)을 통해 $\psi \in W_0$ 가 존재하여, ψ 의 integer translate $\psi(\cdot - k)$ 가 W_0 의 직교정규 기저를 형성함을 알 수 있다. 여기서, 식(2.8)에 의해 $\psi_{jk}(x) = 2^{j/2}\psi(2^j x - k)$ 는 W_j 의 직교 정규 기저를 형성함을 알 수 있다. 그러므로, 식(2.9)(2.10)에 의해 $\{\psi_{jk}, j, k \in Z\}$ 는 $L^2(R)$ 의 직교 정규 기저를 형성함을 알 수 있다. 또한 $\phi_{jk}(x) = 2^{j/2}\phi(2^j x - k)$ 는 조건(2.2)과 (2.5)에 의해 ϕ_{jk} 는 V_j 의 직교정규 기저를 형성한다. 따라서 wavelet 변환은 가장 낮은 해상도(V_j 공간)에서의 근사함수와 차이 공간(W_j)의 근사함수를 유지함으로써 원래의 함수를 주어진 정확도에 따라 근사할 수 있는 다중해상도 분석이라 볼 수 있다. 또한 차이 함수 공간의 합으로 나누는 것은 일종의 다중주파수(subband)분할로 볼 수 있으며 그림2의 QMF(quadrature mirror filter)역시 wavelet 변환의 한 경우이다[Mall 89a].

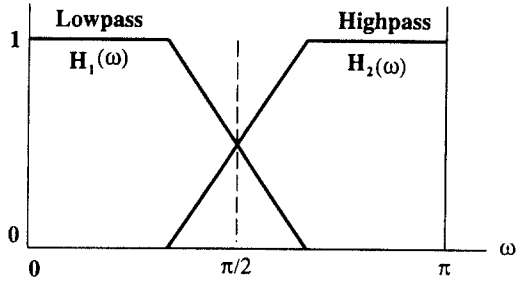


그림 2. 주파수 영역에서의 분할 : QMF[Mall 89a]

$f = (f_1, f_2, \dots, f_N)$, $N = 2^J$ 을 다중 해상도 분석한 결과는 다음과 같다.

$$f = f^\theta + f^{(0)} + f^1 + \dots + f^{J-1},$$

where $f^\theta = 2^{-J} \left(\int_{-x}^x f(x)\phi(x) dx \right) \phi(x)$

$$f^j = 2^{-J} \sum_{k=0}^x \left(\int_x^x f(x)\psi_{jk}(x) dx \right) \psi_{jk}(x)$$

$$= 2^{j-J} \sum_{k=0}^x \left(\int_x^x f(x)\psi(2^j x - k) dx \right)$$

$$\psi(2^j x - k)$$

[예] $f = (10, 4, 8, 2)$, $N = 4 = 2^2 = 2^J$ 일 때, Haar 기저를 이용한 다중 해상도 분석 결과는 다음과 같다.

$$f = (10, 4, 8, 2) = f^\theta + f^0 + f^1$$

$$= 6 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + 1 \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} + 3 \begin{bmatrix} 1 \\ -1 \\ 0 \\ 0 \end{bmatrix} + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

where

$$6 = 2^{-2} \int_0^1 f(x)\phi(x) dx = \frac{1}{4} (10 + 4 + 8 + 2)$$

$$1 = 2^{-2} \int_0^1 f(x)\psi(x) dx = \frac{1}{4} (10 + 4 - 8 - 2)$$

$$3 = 2^{-1} \int_0^1 f(x)\psi(2x) dx = \frac{1}{4} (10 - 4 + 0 + 0)$$

$$3 = 2^{-1} \int_0^1 f(x)\psi(2x-1) dx = \frac{1}{4} (0 + 0 + 8 - 2)$$

3.3 Biorthogonal wavelet 변환

Biorthogonal 표현방법은 orthogonal 표현의 확장된 것으로 Orthogonality를 완화하고 대신 신호를 잘 표현할 수 있는 기저를 설계하는 데 목적이 있다. Biorthogonal 표현이 orthogonal 표현과 다른 점은 O_m 과 V_m 간에 직교성이 존재하지 않는다는 점이며, 따라서 신호를 분해하기 위해 O_m 과 V_m 를 쓸 수 없으며 새로운 딸림 공간 $\tilde{V}_m \in L^2(R)$ 을 정의한다.

$$\dots V_{-2} \subset V_{-1} \subset V_0 \subset V_1 \subset V_2 \subset \dots \quad (11)$$

$$\dots \tilde{V}_{-2} \subset \tilde{V}_{-1} \subset \tilde{V}_0 \subset \tilde{V}_1 \subset \tilde{V}_2 \subset \dots \quad (12)$$

such that $\bigcap_{j \in Z} V_j = 0, \bigcup_{j \in Z} V_j = L^2(R)$

where $V_m \rightarrow L^2(R)$ for $m \rightarrow \infty$,

$$\bigcap_{j \in Z} \tilde{V}_j = 0, \bigcup_{j \in Z} \tilde{V}_j = L^2(R)$$

where $V_m \rightarrow L^2(R)$ for $m \rightarrow \infty$.

또한 각 공간은 낮은 해상도의 공간으로 분해될 수 있다.

$$V_j \oplus W_j = V_{j+1}$$

$$\tilde{V}_j \oplus \tilde{W}_j = \tilde{V}_{j+1}$$

$$V_j \perp W_j$$

$$\tilde{V}_j \perp \tilde{W}_j$$

하지만 두 나누어진 공간은 직교하지 않는다. 따라서 재복원을 위해 딸림 공간 \tilde{V}_j 과 \tilde{W}_j 이 다음 조건(biorthogonality)을 만족하도록 하게된다.

$$\tilde{W}_j \perp V_j$$

$$W_j \perp \tilde{V}_j$$

해상도 공간 $V_m, W_m, \tilde{V}_m, \tilde{W}_m$ 은 각기 wavelet 함수 $\phi(t), \psi(t), \tilde{\phi}(t), \tilde{\psi}(t)$ 의 traslation과 dilation에 의해 완전하게 정의될 수 있으며, 신호의 Biorthogonal분해는 다음과 같다.

$$P_j f = \sum_{k \in \mathbb{Z}} \phi_{jk} \langle \tilde{\phi}_{jk}, f \rangle$$

따라서 신호의 Biorthogonal 다중해상도 표현은 Biorthogonal wavelet 변환과 동일하다.

$$\phi_{m+1}(t) = \phi_m(t) + r_m(t)$$

$$\phi_m(t) = \sum_{n \in \mathbb{Z}} \phi_{mn} \langle \tilde{\phi}_{mn}, f \rangle$$

$$r_m(t) = \sum_{n \in \mathbb{Z}} \psi_{mn} \langle \tilde{\psi}_{mn}, f \rangle$$

IV. Packet Wavelet변환

영상 압축에서는 에너지 집약과 인간 시각에 맞는 압축이 요구되며 동영상 압축에서 나타나는 예측오차를 압축하기 위해서는 랜덤신호에서 edge성분을 뽑아내는 것이 중요하며 edge 성분은 어느 일정부분에 집중될 수 있다. 따라서 동영상 압축에서 나타나는 예측오차압축과 같은 문제에서는 4개 주파수 대역으로 나누는 것이 최선이 아니다. 최근 wavelet이론의 일반화에 대한 연구가 진행 중이며 wavelet변환이 가진 두개의 매개 변수인 scale과 position에 frequency매개변수를 추가한 packet wavelet이라는 개념이 발표되었다 [Coif 92]. scale 매개변수는 wavelet변환에서와 같이

phase space에서의 윈도우의 크기와 형태를 결정짓게 되며 나머지 매개변수 position과 frequency를 이용하여 phase space내의 주어진 위치에 윈도우를 배치할 수 있다.

Packet wavelet변환의 한 예를 그림3에서 다른 변환 방법과 비교할 수 있다. wavelet변환에서는 scale과 position이 정해지면 즉 기저가 정해지면 phase space tiling이 결정되지만 packet wavelet변환을 이용하게 되면 frequency매개변수를 바꿈으로써 무한대의 tiling배치상태가 가능하다. 따라서 packet wavelet변환에서의 문제는 신호를 가장 표현할 수 있도록 무한대의 가능한 tiling배치로부터 주어진 기준에 맞는 phase space tiling배치를 선택하는 방법이 문제가 되며 일종의 최적화문제로 볼 수 있다.

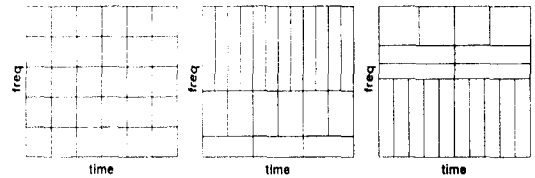


그림 3. Phase space tiling 비교 : (a) Windowed Fourier변환 (b) Wavelet 변환 (c) Packet Wavelet 변환 [Ebra 92]

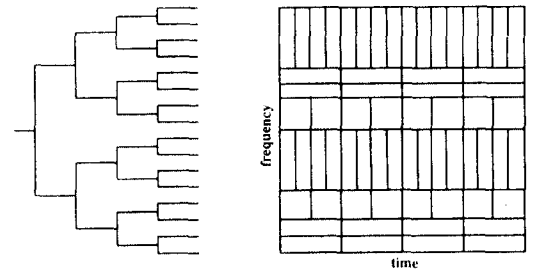


그림 4. Packet Wavelet변환의 Binary tree decomposition [Ebra 92]

그림4에서는 Packet wavelet변환과 Adaptive Sub-band Decomposition과 동일함을 볼 수 있다. 그림에서 굵은 선은 서브밴드분해된 영역을 나타낸다. windowed fourier변환의 경우 full tree에 해당되는 binary tree가 된다. 영상처리를 위해서는 고차원 packet wavelet변

환이 필요하며 분해가능한 필터를 여러 번 적용시킴으로써 원하는 차원의 필터를 구성할 수 있다. 그림5에서는 영상처리를 위한 2차원 서브밴드 분해방법을 도시하고 있다.

Packet wavelet변환은 주어진 응용에 맞는 기준에 의해 가장 최선의 phase space tiling배치를 찾는 것이 핵심 문제이며 따라서 원하는 응용에 맞게 기준을 정하는 것이 큰 문제가 된다. 실제로 영상압축의 응용에서는 엔트로피를 최소화하는 것이 목적이므로 엔트로피를 최소화하는 phase space tiling배치를 찾게 된다 [Ebra 92]. 최근에는 정해진 비트율에 맞추기위해 rate-distortion의 합성식의 값을 최소화하는 방법이 연구되고 있다.

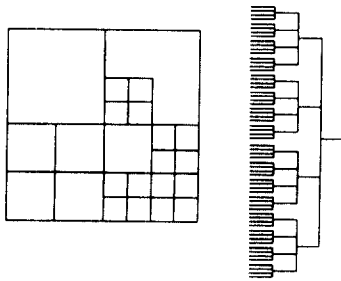


그림 5. 2D Packet Wavelet변환의 Quadtree decomposition [Ebra 92]

4.1 다차원신호에의 확장

다중해상도 표현과 wavelet 변환은 다차원으로 확장할 수 있으며, 여기서는 단순히 2차원에 대해 국한한다. 2차원 신호 분해는 다음과 같은 식으로 이루어진다.

$$V_{j+1} = V_j \oplus W_j$$

$$W_{j+1} = W_j^{(H)} \oplus W_j^{(V)} \oplus W_j^{(D)}$$

역시 각각의 해상도 공간은 다음과 같은 separable wavelet 기저에 의해 span 가능하다.

$$\phi^{(H)}(k_1, k_2) = \phi(k_1) \times \phi(k_2)$$

$$\psi^{(H)}(k_1, k_2) = \psi(k_1) \times \phi(k_2)$$

$$\psi^{(V)}(k_1, k_2) = \phi(k_1) \times \psi(k_2)$$

$$\psi^{(D)}(k_1, k_2) = \psi(k_1) \times \psi(k_2)$$

V. Wavelet변환을 이용한 동화상압축기법

Wavelet변환을 이용한 영상압축기법은 주로 정영상을 위주로 제시되었으며, 동화상을 위주로 한 것은 소수이며 중복성 제거의 순서와 3차원변화의 사용여부에 따라 기존의 연구를 분류하고자 한다[Barn 94].

5.1 H.261의 변형

가장 쉽게 생각할 수 있는 wavelet변환을 이용한 동화상압축기법은 ISDN용 비데오폰의 표준안으로 제시된 H.261 동화상압축기법에서 DCT와 DCT와 관계된 블럭들(예를 들어 양자화, 엔트로피코딩)을 wavelet변환을 기반으로 새로 설계하는 것이다. 그림6에서 보는 바와 같이, 우선 기준 프레임(reference frame)에 대하여 wavelet변환을 적용하여 전송하고, 내부의 decoder loop를 통하여 복원된 기준 프레임과 다른 프레임 간에 움직임 추정을 통해 시간 상관성(temporal re-

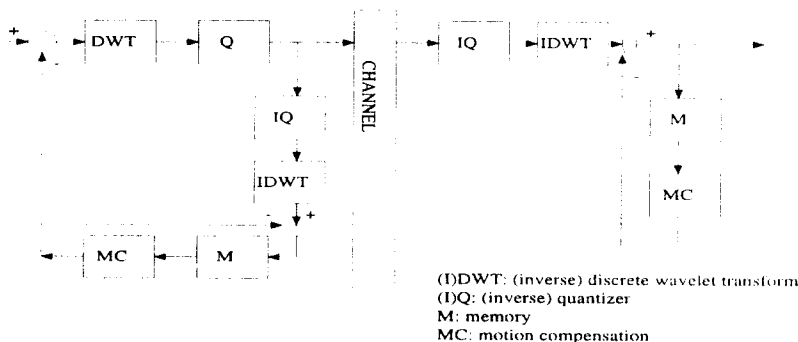


그림 6. 동화상 압축기법(1) : temporal coding followed by spatial coding

dundancy)를 제거하고, 움직임 오차에 대하여 wavelet 변환을 적용하는 방법이다. 이 방법은 multiresolution 기법을 적용할 수 있으며, 즉 decoder가 memory도 적고, 계산속도도 느릴 경우, 낮은 해상도의 영상을 보냄으로써 다양한 사용자에게 서비스를 제공할 수 있다는 장점이 있다. 또한, packet망을 통해 영상을 전송할 경우, 저주파영상신호에 높은 우선순위를 줌으로써 전송선의 불량에도 좋은 영상품질을 보장할 수 있다.

5.2 대역별 움직임 추정기법

그림7에서 보여지는 영상압축기법의 경우, wavelet 변환을 가친 후 얻어지는 프레임 간의 대응하는 대역 영상간에 움직임 추정을 통하여 상관성을 제거하는 방법이다. Zhang의 연구가 대표적인데, 구체적인 방법은 아래 그림 8(a)에 제공되고 있다. 즉 저대역 영

상간에 움직임 추정이 행해지고, 높은 주파수 대역의 영상들끼리는 상관성이 존재하므로 해당주파수 대역에서 추정된 움직임 정보는 다음 주파수 대역에서의 움직임 추정의 초기치로 쓰일 수 있다. sparial domain에서의 spatial decomposition과 같이 wavelet 변환된 영상에 대해서도 같은 기법을 적용할 수 있는데 그림 8(b)에서는 고정된 크기의 블록이 아닌 가변크기의 블록에 대해 움직임 추정을 행하는 변형된 방법이며, 부가정보가 필요하지만 대체로 좋은 성능을 보인다. 낮은 대역간의 움직임 정보를 높은 대역의 움직임 정보의 초기치로 이용하는 것은 계층 부호화기법과 일치하며, 원영상에 대한 움직임 추정성능보다는 떨어지지만 전체적인 움직임정보의 감소가 이루어지며, 이웃되는 대역들에서의 움직임 정보를 서로 이용할 경우 움직임 추정성능은 거의 최적에 근사함이 밝혀 있다[Vial 92].

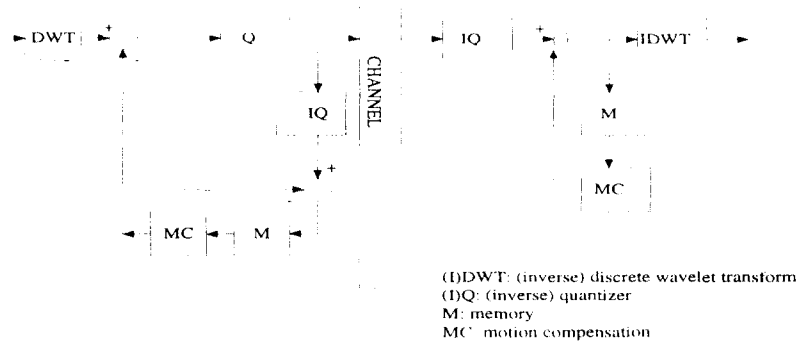


그림 7. 동화상 압축기법(2) : spatial coding followed by temporal coding

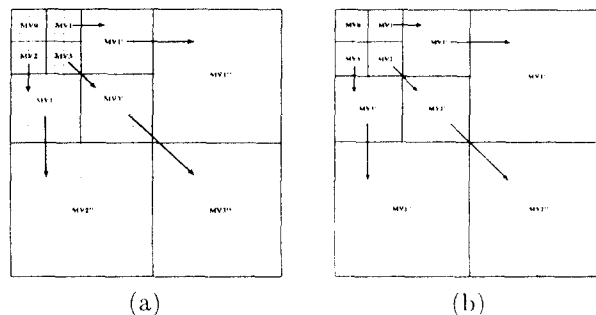


그림 8. (a) Zhang의 고정 크기 블록 움직임 추정 및 예측
 (b) 변형된 가변 크기 블록 움직임 추정 및 예측

이 방법은 wavelet 변환 후 각 대역에 대하여 각각 영상압축이 이루어지므로, multiresolution/multifrequency 성질 모두를 만족하며, 영상의 점진적 전송(progressive transmission) 등에 적절한 기법이다. 성능비교에 따르면, H.261 변형 방법은 low bit rate 부호화 기법에, 대역별 움직임 추정기법을 이용한 방법은 high bit rate에 적절한 기법임이 밝혀졌다.

5.3 계층적 영상부호화 기법

그림9에 제시된 동화상압축기법은 앞에 제시된 방법에 비해 복잡하며, 고해상도 영상이 필요할 경우, 이미 전송된 대역의 양자화오차가 다시 양자화되어서 전송된다.

5.4 3차원 변환 영상부호화기법

2차원 wavelet 변환을 바로 확장하여 3차원 wavelet 변환을 그림10 처럼 공간축과 시간축에 대해 적용할

수 있다. 하지만 이 방법은 다음과 같은 세 가지 단점을 지님으로써 잘 쓰여지지 않는다.

우선 첫째로, 여러 개의 연속된 프레임에 대하여 부호화를 적용함으로써, 지연시간이 길어지게 된다. 이는 비디오폰과 같은 대화형 응용의 경우에는 적용할 수 없다. 또한 여러 프레임을 동시에 메모리에 유지해야하는 비용문제도 무시할 수 없다. 두번째로, 공간상에서 보여지는 상관성과는 달리, 프레임간에 존재하는 상관성은 하나의 물체에 대한 영상이 아니게 되며, 따라서 물체의 움직임, 새로운 물체의 출현, 최악의 경우, scene change의 경우, 3차원 변환 부호화는 화질이 저하되게 된다. 이러한 문제에 대한 해결책으로는 프레임 간에 움직임 보정을 기친 후에 3차원 변환을 적용하는 방법이 있을 수 있다. 세번째의 단점은 시간축에 대해 2 tap 이상의 filtering을 적용할 때, 벗어나는 영역에 대한 신호확장을 정의하기가 곤란하다는 점이다. 해결책의 하나는 시간축에 대해 프레임을

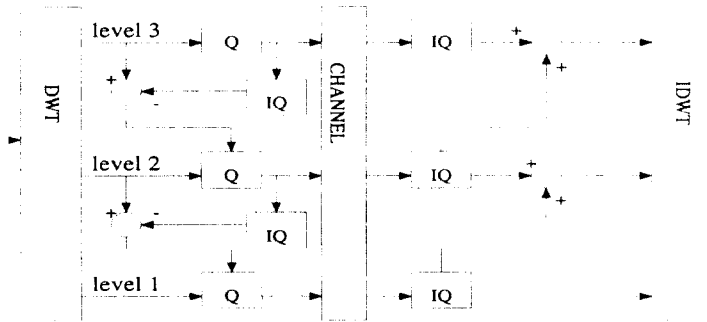


그림 9. 동화상 압축기법(3) : hierarchical coding with requantization of the subbands

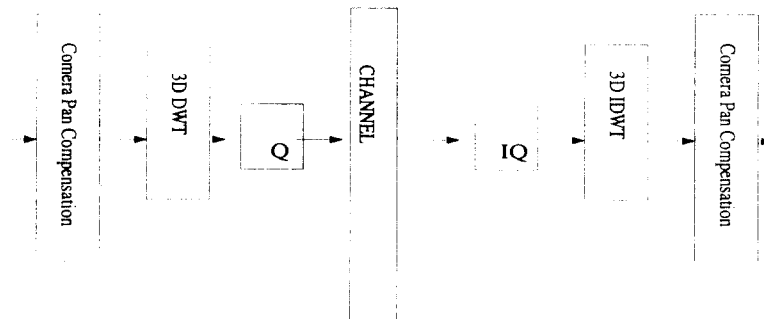


그림 10. 동화상 압축기법(4) : 3 dimensional wavelet transform coding with camera pan compensation

grouping하지 않고 계속되는 프레임에 대해 filtering을 적용하는 것이지만 이렇게 될 경우, MPEG-1에서 제공하는 fast forward나 refresh mode를 지원할 수 없다는 점이다. 이러한 문제들 때문에 3차원 변환시 시간축에 대해서는 신호확장(signal extension)이 필요 없는 Haar filtering을 적용하지만 Haar를 적용할 경우 frame differencing기법과 거의 동일하므로 과연 변환을 적용한다고 볼 수 있느냐는 문제도 있다.

VI. 제안된 동화상 압축기법

여기서는 전 절에서 설명된 동화상 압축기법 중 대역별 움직임 추정을 변형한 방법을 소개하기로 한다 [Kim 93]. 전체적인 부호화기는 다음 그림11과 같으

며, CBR(constant bit rate)를 위한 rate control부분은 제시되지 못했다. 먼저 입력된 프레임은 기준프레임의 경우, wavelet변환과 엔트로피 코딩을 거쳐서 전송된다. 이제 내부의 decoder loop를 통하여, 기준 프레임이 복원된다. 동시에 기준 프레임과 임의의 프레임에 대해, wavelet변환 후의 각 대역별로 움직임 추정이 이루어진다. 추정된 움직임 벡터는 decoder loop를 가진 복원영상에 적용이 되고, 움직임 오차는 다시 wavelet변환을 통해 부호화된다.

다음 절에서는 부호화기의 각 부분에 대하여 핵심적인 사항만 설명하기로 한다.

6.1 wavelet변환부

우선 wavelet변환을 하기 위해, 여기서는 Daubechie

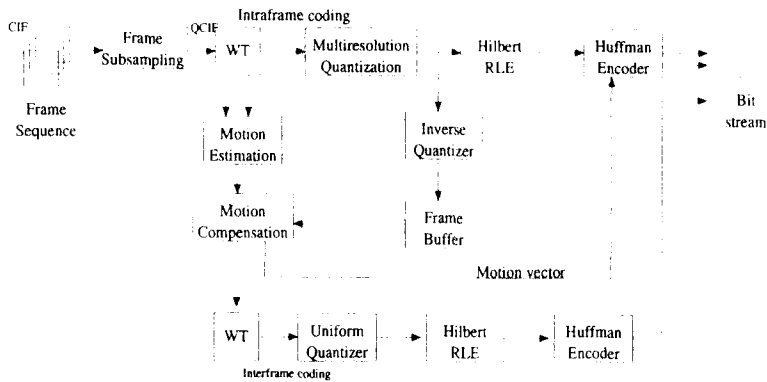


그림 11. Video Coder

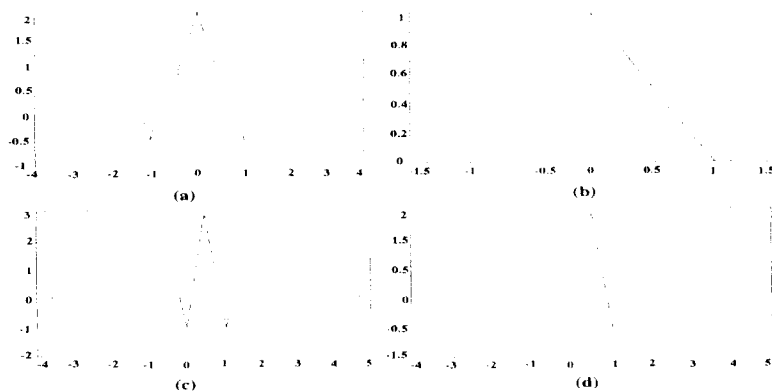


그림 12. Scaling functions $\tilde{\phi}$, $\tilde{\psi}$ and wavelets ψ , $\tilde{\psi}$ (a) Scaling functions $\tilde{\phi}$, (b) Scaling functions $\tilde{\phi}$, (c) Wavelet ψ , (d) Wavelet $\tilde{\psi}$

가 제안된 biorthogonal 9-3 필터를 사용한다. 그림12는 9-3필터의 모양을 보이고 있으며, 표5는 필터계수이다.

표 5. Filter coefficients for the Daubechie 9-3 filter

n	0	± 1	± 2	± 3	± 4
$2^{-1/2} h_n$	45/64	19/64	-1/8	-3/64	3/128
$2^{-1/2} \tilde{h}_n$	1/2	1/4	0	0	0

그림13의 동화상압축에서 표준영상으로 쓰이는 miss-america 영상의 임의의 한 프레임에 대한 wavelet변환 결과이다. 변환결과 중 화소값이 3이상되는 것은 하얀 점으로 표기하였으며, 전체적으로 극히 수가 제한됨을 알 수 있다. 따라서 RLE(run-length-encoding)의 부호화모델을 이용할 경우, 0의 이어지는 길이를 최대화할 수 있어야 하며, 제시된 모델에서는 출단위의 RLE대신 해당구역 내에 가장 많이 머무르는 프랙탈 커브를 이용하여 RLE를 수행한다. 다음 절에서 자세한 사항을 보이기로 한다.

6.2 스칼라양자화와 Hilbert스캔을 통한 wavelet 변환계수의 부호화

wavelet변환을 거친 영상은 각 대역별로 비트를 할당받아 고주파대역은 인간시각이 변화율에 둔감하기 때문에 양자화 간격을 크게 하고, 낮은 주파수의 경우에 양자화 간격을 줄임으로써 전체 화질을 떨어뜨리지 않아야하며, 여기서 제안된 양자화기는 최적양자화식을 풀지않고, 인간시각에 맞는 양자화 테이블(그림14)를 이용하여 양자화시켰다.

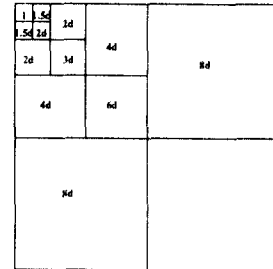


그림 14. Bit Rate for each Frequency Band

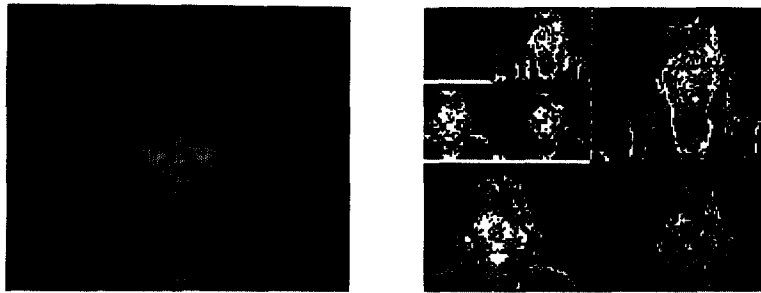


그림 13. (a) Sample image (b) Wavelet Transform of Sample image (white points means the values above 3)

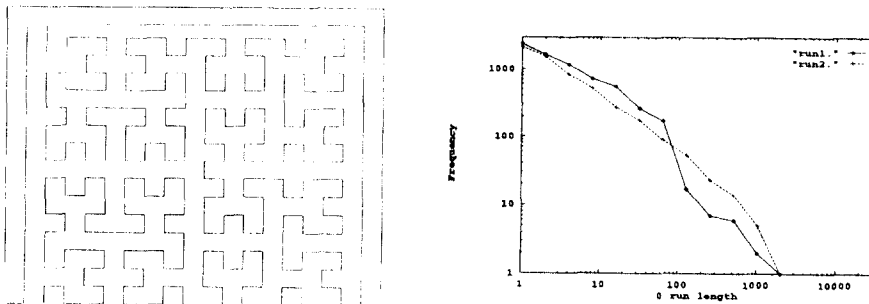


그림 15. Hilbert curve scanning(a) and its performance(b)

선 절에서 제시되었던 wavelet변환의 결과를 보면 고주파 영상의 경우, 대부분이 0의 값을 갖는 것을 알 수 있으며, 따라서 문제는 이러한 많은 0을 어떻게 효과적으로 부호화하는냐의 문제이며, 이는 주어진 영역에 최대한 머물러 있을 수 있는 프랙탈 커브의 일종인 Hilbert curve를 이용하여, 효과적으로 압축할 수 있다. 그림15는 줄단위로 RLE(run length encoding)을 적용시킨 결과와 Hilbert 커브를 이용하여 RLE를 시킨 결과이다.

6.3 대역간 움직임 추정

전 상에서 제시된 바와 같이 각 대역별로 움직임 추정을 행하게 되며, 낮은 주파수 영역에서의 움직임 추정이 이후 주파수에서의 움직임 추정의 초기치로 사용되므로 반화소단위의 움직임 추정을 행하였다. 그림16은 개략적인 움직임 추정기법을 보여주고 있으며, 그림에서 보여지듯이 저대역영상에 대해서는 가로11, 세로9의 클럭크기를 이용하며, 중대역에서의 움직임 추정은 이후의 움직임추정의 초기치로 사용되므로 반화소 단위로 움직임 추정을 정확히 하였다.

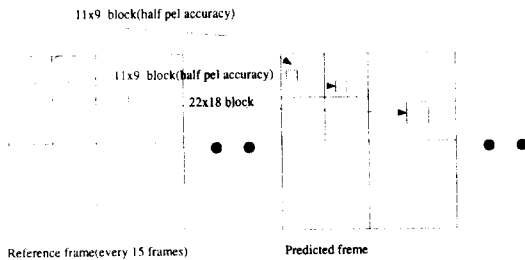


그림 16. Multifrequency motion estimation scheme

6.4 움직임 추정성능

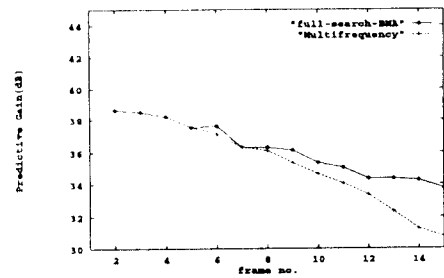
움직임은 실제 이루어진 움직임을 정확히 예측하느냐로 평가해야하지만 정확한 평가척도가 없으므로, 다음과 같은 두 가지 개관적인 평가척도를 이용하여 평가하게 되며, 그림17는 miss-america 10프레임에 대한 움직임 추정성능을 도시하였다. 표를 보면 pre-

dictive gain면에서는 성능이 떨어지지만 entropy측면에서는 성능이 증가하는 것을 알 수 있다.

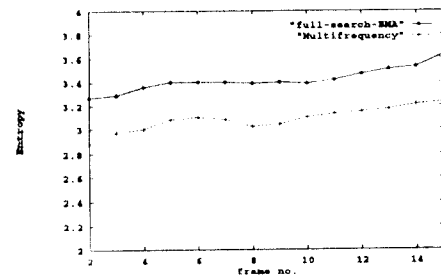
하지만 대역간 움직임 추정을 사용하는 이유는 점진적 전송이 가능하다는 장점외에 계산량이 준다는 장점이 있으며, 이는 초저속 영상통신기가 성공하기 위한 좋은 변이 된다. 다음은 일반적으로 많이 쓰이고 있는 전역탐색(full search)방법과 3단계 탐색방법(3 step search), 그리고 제안된 대역 움직임 추정기법을 비교하고 있으며, 성능에 비해, 낮은 계산량을 보이고 있다.

$$PG(dB) = 10 \log_{10} \frac{255^2}{\sigma_{BFD}^2}$$

$$Entropy = - \sum_{i=1}^{255} P_i \log_2 P_i$$



(a)



(b)

그림 17. (a) Predictive gain for Miss America sequence: full search BMA and multifrequency motion estimation (b) Entropy for Miss America sequence : bmafull : full search BMA and multifrequency motion estimation

표 6. Time complexity between different motion estimation techniques : R : search range, n : block size

Algorithm	search points/block	average search point/block
Full Search	$(R+1)^2$	81
Three Step Search	27	27
Multifrequency	$(R1+1)^2$ /low pass, $(R2+1)^2$ /high pass	51

다음 그림18는 miss-america 10프레임에 대하여 압축율과 PSNR을 보인 것이며, 현재 75대 1정도에서 33dB정도의 성능을 보이고 있으며, 인간시각이 가지고 있는 temporal masking effect로 주관적 화질은 크게 손상이 없다.

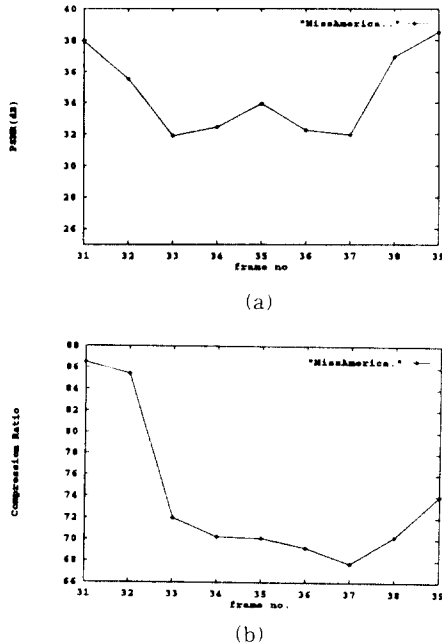


그림 18. PSNR for Miss America sequence (b) Compression ratio for Miss America sequence

Ⅶ. 맺음말

본 기고에서는 초저속영상 부호화기법을 표준화하기 위한 ITU의 활동을 중심으로, 초저속 영상통신의 응용분야, 요구사항, 그리고 현재 진행되고 있는 연구들, 특히 단기표준안이 되고 있는 DCT기반의 TMN2와 장기적인 표준안으로 제시되고 있는 프랙탈코딩, 모델베이스 코딩의 간략한 설명을 덧붙였다.

이어서, wavelet변환을 기반으로 한 동화상압축기법들을 4개의 분류안에서 분류하고 각각을 요구사항 측면에서 분석하였다. 또한 구체적으로 대역간의 움직임 추정을 통한 동화상 압축기법을 제안하고, 세물레이션하여 평가하였다. 제안된 부호화기는 요구사항 측면에서 평가되었다. 하지만 제안된 기법은 고정

비트율에 데이터량을 맞추는 buffer control 부분이 추가되어야하며, 색조 영상에 대해서는 고려하고 있지 못하다.

Wavelet변환을 기반으로 한 동화상압축기법은 현재 단기안으로는 고려되고 있지 않지만 DCT로는 화질과 비트율이 제한이 되므로 반드시 전환이 있을 것으로 생각하고 있다.

참 고 문 헌

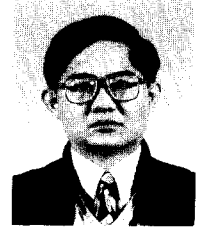
- [Anto 92] Marc Antonini, Michel Barlaud, and et.el., "Image Coding Using Wavelet Transform," *IEEE Transactions on Image Processing*, 1(2), April 1992.
- [Barn 94] H.Barnaud, "Image and Video Coding Using a Wavelet Decomposition," *PhD thesis. Delft Technische Universiteit*, 1994
- [Bier 86] Matthias Bierling and Robert Thoma, "Motion Compensating Field Interpolation Using a Hierarchically Structured Displacement Estimator," *Signal Processing*, 11, 1986.
- [Coif 92] Ronald R. Coifman and Mladen Victor Wickerhauser, "Best-Adapted Wave Packet Bases," *in preprints*.
- [Comb 89] J.M. Combes, A. Grossman, and Ph. Tchamitchian(ed), "Wavelets: Time Frequency Methods and Phase Space(2/e)," *Springer-Verlag*, 1989.
- [COST 93] COST 211ter, "Simulation model for very low bite rate image coding(SIM3)," *NTR, SIM(93)24*, 1993.
- [Daub 88a] Ingrid Daubechies, "Othonormal Bases of Compactly Supported Wavelets," *Communications on Pure and Applied Mathematics*, XLI, 1988.
- [Daub 88b] Ingrid Daubechies, "Time-Frequency Localization Operators: A Geometric Phase Space Approach," *IEEE Trans. on Information Theory*, 38(4), July 1988.
- [Daub 90] Ingrid Daubechies, "The Wavelet Transform, Time-Frequency Localization and Signal Analysis," *IEEE Transactions on Information Theory*, 36(5), Sep. 1990.

- [Devo 92] Ronald A. Devore, Bjorn Jawerth and Bradley J. Lucier, "Image Compression Through Wavelet Transform Coding," *IEEE Transactions on Information Theory*, **38** (2), March 1992.
- [Ebra 92] Touradj Ebrahimi, "Perceptually Derived Localized Linear Operators: Application to Image Sequence Compression," *Ecole Polytechnique Federale De Lausanne DThesis*, 1992.
- [Feau 91] J. C. Feauveau, P. Mathieu, M. Barlaud, M. Antonini, "Recursive Biorthogonal Wavelet Transform for Image Coding," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing, IEEE*, 1991.
- [Gall 91] D.L.Gall, "MPEG: A Video Compression Standards for Multimedia Applications," *Comm. of the ACM*, **34**(4), April 1991.
- [Jacq 92] A. Jacquin, "Image coding based on a fractal theory of iterated contractive image transformations," *IEEE Transactions on Image Processing*, **1**(1), Jan. 1992.
- [Jain 91] J. R. Jain and A. K. Jain, "Displacement Measurement and Its Application in Inter-frame Image Coding," *IEEE Trans. on Communication*, **COM-29**(12), Dec. 1991.
- [Kim 93] Seong-Whan Kim and Heung-Kyu Lee, "Video Coding with Wavelet Transform on the Very Low Bit Rate Communication Channel," *IS&T/SPIE Symposium on Electronic Imaging Science & Technology*, 1993.
- [Anto 92] 한국통신 연구개발단, "초저속 전송을 위한 영상 통신 기술 워크숍 93," 한국통신, 1993.
- [ktm 93] 한국통신 연구개발단, "10 Kbps급 동영상 전송을 위한 모뎀기반 부호화연구," 명지대학교, 1993.
- [Lewi 90] A.S.Lewis and G. Knowles, "Video Compression Using Wavelet Transform," *Electronics Letters*, **26**(2), March 1990.
- [Liou 91] M.Liou, "Overview of the P×64 Kbit/s Video Coding Standards," *Comm. of the ACM*, **34**(3), April 1991.
- [Mall 89a] Stephane G. Mallat, "Multi-Frequency Channel Decomposition of Images and Wavelet Models," *IEEE Trans. on Information Theory*, **11**(7), July 1989.
- [Mall 89b] Stephane G. Mallat, "A Theory for Multiresolution Signal Decomposition: The Wavelet Representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**(7), July 1989.
- [Meri 85] K. Metin Uz, Martin Vitterli and Didier J. LeGall, "Interpolative Multiresolution Coding of Advanced Television with Compatible Subchannels," *IEEE Trans. on Circuits and Systems for Video Technology*, **1**(1), March 1991.
- [Musm 89] H.Musmaann, M.Hotter, and J.Ostermann, "Object Oriented Analysis-Synthesis Coding of Moving Images," *Signal Processing: Image Communication*, **1**(1), 1989.
- [Ohta 85] Mutsumi Ohta, Mitsuharu Yano, and Takao Nishitani, "Entropy Coding for Wavelet Transform of Image- And Its Application for Motion Picture Coding," *SPIE Visual Communications and Image Processing*, **1605**, 1991.
- [Stra 89] Gilbert Strang, "Wavelet and Dilation equations: A brief introduction," *SIAM Review*, 1989.
- [Vetr 89] Martin Vetterli, "Multiresolution Coding Techniques for Digital Television: A Review," *Columbia Univ.*, 1989.
- [Vial 92] J. Vial, M. Pecot, P. Tourtier, and Y. Thomas, "In-Band Interpolation Applied to Motion-Compensated Subband Coding," *Signal Processing of HDTV III*, Elsevier, 1992.
- [Wood 69] R.C. Wood, "An optimum quantization," *IEEE Trans. Inform.Theory*, **vol.IT-15**, March 1969.
- [Wood 91] J.W. Woods(ed), "Subband Transform Coding," *Kluwer-Academic*, 1991.
- [Zhan 91] Ya Qin Zhang and Sohail Zafer, "Motion-Compensated Wavelet Transform Coding



김 성 환

- 1969년생
- 1991년 : 동국대학교 전산학과 졸업
- 1993년 : 한국과학기술원 전산학과 석사졸업
- 1993년 ~ 1994년 현재 : 한국과학기술원 전산학과 박사과정
- 주관심분야 : wavelet 변환, image and video coding



이 흥 규

- 1955년생
- 1978년 : 서울대학교 전자공학과 졸업
- 1981년 : 한국과학기술원 전산학과 석사졸업
- 1984년 : 한국과학기술원 전산학과 박사졸업
- 1985년 ~ 1986년 : 미시간 대학교 연구원
- 1986년 ~ 현재 : 한국과학기술원 전산학과 부교수
- 주관심분야 : 영상처리, VLSI, 실시간 및 고장허용 처리