

論文2002-39SP-2-2

절대값 대칭성과 사전 분류를 이용한 고속 Matching Pursuit (Fast Matching Pursuit Using Absolute Symmetry and Classified Dictionary)

吳錫炳*, 全炳宇*

(Seok Byoung Oh and Byeung Woo Jeon)

요약

비디오 데이터 부호화를 위한 여러 가지 방법들 중 matching pursuit을 이용한 비디오 부호화 방법은 저속 채널상에서 뛰어난 효율을 보이고 있으나, 기존의 블록 기반 변환 부호화 방법에 비하여 연산량이 월등히 많기 때문에 비디오 부호화에 적용하는 데는 어려운 점이 있다. 본 논문에서는 matching pursuit의 전체 계산량 중 가장 큰 비중을 차지하는 내적 연산량을 줄이기 위해 matching pursuit의 내적연산에 사용되는 1차원 사전(Dictionary)을 구성하는 각각의 기본 신호들의 절대값이 대칭성을 가진다는 특성과, 크게 화질에 영향을 주지는 않지만 연산량을 줄이는 데는 도움을 줄 수 있도록 변형된 사전을 사용할 수 있다는 것, 그리고 주파수 특성을 이용하여 2차원 사전이 미리 분류될 수 있다는 사실을 이용한다. 실험결과에 나타난 바와 같이 제안하는 방법은 큰 화질 열화 없이, Neff가 제안한 사전의 분리성(Separability)을 이용한 기존의 고속 방법보다 내적 연산량에 소요되는 곱셈량을 약 1/8로 감소시키는 효과를 가져온다.

Abstract

Although the matching pursuit is effective for video coding at low bit rate, it has a problem since it needs much more calculation than the conventional block-based video coding method. The proposed fast matching pursuit method reduces inner product calculation that takes the most part of entire calculation by utilizing the symmetry of the absolute values of the one-dimensional Gabor dictionary bases, the modified dictionary which allows faster matching without causing image quality degradation, and a property of the two-dimensional Gabor dictionary that can be grouped in advance to four classes according to its frequency characteristics. Proposed method needs only about 1/8 of multiplications compared to the well-known separability-based fast method proposed by Neff.

Key Word : matching pursuit, Gabor, low bit rate video

I. 서론

Matching pursuit은 신호를 선형적으로 분해하여 주

* 正會員, 成均館大學校 電氣電子 컴퓨터工學部

※ 본 논문은 정보통신부에서 지원하는 1999년도 대학 기초 연구 지원 사업으로 수행되었음.

接受日字:2000年4月4日, 수정완료일:2002年2月15日

어진 신호의 특성을 파악하고자 하는 목적으로 Mallat에 의해 최초로 제안된 방법이다^[1]. 이 방법은 신호를 분해함에 있어서 windowed Fourier 변환이나 웨이블릿 변환처럼 시간영역과 주파수영역을 동시에 고려하여 원 신호와 가장 비슷한 성분을 가지는 기본 신호들을 에너지의 크기 순으로 찾아내어 전체적인 원 신호의 특성을 파악하는 이론이다. 이러한 1차원 신호의 특성을 분석하는 방법을 2차원 영상신호에 시도한 결과^[2-4], 특히 저속 채널상에서의 움직임 보상 오차신호 부호화

에 있어 기존의 블록기반 변환 부호화 방법보다 PSNR 이나 주관적인 화질면에서 모두 뛰어난 개선효과가 있다는 것이 밝혀졌다^[4]. 그 이유는, 기존의 블록기반 방법이 고정된 위치에서만 변환 과정을 수행하는데 반하여 matching pursuit은 고정된 블록구조와 상관없이 자유로운 위치에서 변환을 이룰 수 있어 주어진 입력신호에 최적의 변환이 가능하기 때문이다. 또한 적용되는 기본 신호들이 DCT와 같은 기존 커널 함수보다 특히 커널 양쪽 경계면에서 더욱 부드러운 특성을 가지고 있기 때문에 영상의 블록화 현상과 같은 블록 기반 방법의 단점을 극복할 수 있다. 그러나 matching pursuit에 사용되는 기본 신호들은 서로 직교(Orthogonal)하지 않아 하나의 최적 기본 신호를 찾아내기 위하여 많은 수의 기본 신호들에 대한 내적연산 과정을 모두 거쳐야 하기 때문에 계산량이 상당히 많다는 단점 또한 가지고 있다. 이러한 단점을 극복하기 위하여 여러 연구들이 진행되어 왔으나^[4-7], 지금까지 연구되어진 고속 방법으로도 아직 실시간 화상 통신등에 이용하기에는 어려운 점이 있다. 본 논문에서는 화질열화를 거의 주지 않고 계산량을 기존 방법^[4]보다 약 8배정도 추가로 줄일 수 있는 새로운 방법을 제안하고자 한다.

본 논문은 다음과 같이 구성되어 있다. 먼저, 2장에서는 matching pursuit의 기본개념과 이에 소요되는 사전(dictionary)을 만드는 과정을 살펴보고, 3장에서 움직임 보상 오차신호에 적용된 matching pursuit 기반 영상 부호화 과정에 대한 자세한 설명과 기존의 고속 방법 중 제안하는 방법에서도 함께 사용될 2차원 Gabor 사건의 분리성을 이용한 고속 방법에 대해 알아보며, 4장에서는 제안된 방법에 대해서 구체적으로 서술하며, 5장에서 실험결과와 연산량을 비교한 후, 6장에서 결론을 맺는다.

II. Matching Pursuit의 기본개념과 사건의 구성

2.1 사건의 구성

Matching pursuit 이란 주어진 원 신호를 식 (1)과 같은 형태로 분해하는 방법이다. 여기서 g_n 은 n 번째 단계에서 입력신호를 분해함에 있어서, 사전을 구성하는 기본 신호들 중 입력신호의 가장 많은 에너지를 줄일 수 있도록 선택된 기본 신호이고, a_n 은 선택된 기본신호에 곱해지는 확장계수를 나타낸다.

$$f(t) = \sum_{n=0}^{\infty} a_n \cdot g_n(t) \quad \text{식 (1)}$$

사전은 이를 구성하는 방법에 따라 여러 형태가 있을 수 있으나, 가장 잘 알려진 것은 Gabor함수를 이용한 Gabor 사전이다^[4]. 사전을 구성하는 기본 신호들은 하나의 기본 윈도우 함수 $g(t)$ 에 스케일링과 평행이동(translation), 그리고 변조(modulation) 과정을 가함으로써 만들어진다. 기본 윈도우 함수로서는 식 (2)와 같은 가우시안 윈도우 함수가 많이 이용되고 있다.

$$g(t) = \sqrt[4]{2} e^{-\pi t^2} \quad \text{식 (2)}$$

선택된 윈도우 함수에 가하는 스케일링 요소를 s , 평행이동 요소를 u , 변조 요소를 ξ 라고 표현하면, 각 요소들의 조합은 하나의 변수 $\gamma = (s, u, \xi)$ 로 표현되고 이를 통해 만들어지는 1차원 기본 신호들은 식 (3)으로 표현되어진다.

$$g_\gamma(t) = \frac{1}{\sqrt{s}} g\left(\frac{t-u}{s}\right) e^{i\xi t} \quad \text{식 (3)}$$

식 (3)에 대해 다양한 조합 (s, u, ξ)을 적용시킨 함수 $g_\gamma(t)$ 들을 모두 모은 것이 사전이며, 본 논문에서 사용할 사전을 구성하는 기본 신호들에 대한 파라미터 (s, u, ξ)들의 조합이 표 1에 나타나 있다. 표 1의 각 파라미터 값과 가우시안 윈도우 함수를 기본으로 해서 만든 이산시간(discrete-time) 1차원 Gabor 기본 신호들은 식 (4)와 같다. 여기서 각 기본 신호들은 에너지가 1이 되도록 정규화 된다. 식 (4)의 K_γ 는 이를 위한 정규화 상수이며, N 은 각 기본 신호함수의 길이로서, 0이 아닌 구간의 길이를 말한다.

$$g_\gamma[i] = K_\gamma \cdot g\left(\frac{i - \frac{N}{2} + 1}{s}\right) \cdot \cos\left(\frac{2\pi\xi(i - \frac{N}{2} + 1)}{16} + \phi\right) \quad \text{식 (4)}$$

이렇게 구성된 1차원 사전을 이용해서 영상에 적용될 2차원 Gabor 사전을 만드는 방법은 식 (5)와 같다.

표 1. 1차원 Gabor 함수 정의표^[4]
Table 1. Parameters defining the 1-D Gabor dictionary^[4]

Basis				N	비고(4.1절 참조)
1	1.0	0.0	0	1	Pseudo Even
2	3.0	0.0	0	5	Pseudo Even
3	5.0	0.0	0	9	Pseudo Even
4	7.0	0.0	0	11	Pseudo Even
5	9.0	0.0	0	15	Pseudo Even
6	12.0	0.0	0	21	Pseudo Even
7	14.0	0.0	0	23	Pseudo Even
8	17.0	0.0	0	29	Pseudo Even
9	20.0	0.0	0	35	Pseudo Even
10	1.4	1.0	$\pi/2$	3	Pseudo Even
11	5.0	1.0	$\pi/2$	9	Pseudo Odd
12	12.0	1.0	$\pi/2$	21	Pseudo Odd
13	16.0	1.0	$\pi/2$	27	Pseudo Odd
14	20.0	1.0	$\pi/2$	35	Pseudo Odd
15	4.0	2.0	0	7	Pseudo Even
16	4.0	3.0	0	7	Pseudo Even
17	8.0	3.0	0	13	Pseudo Even
18	4.0	4.0	0	7	Pseudo Even
19	4.0	2.0	$\pi/4$	7	N/A
20	4.0	4.0	$\pi/4$	7	N/A

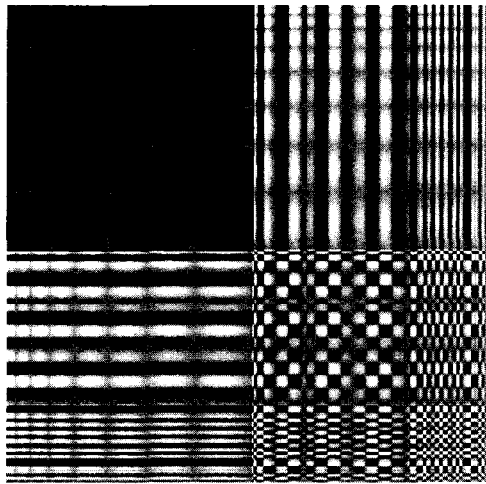


그림 1. 2차원 Gabor 사전
Fig. 1. 2-D Gabor dictionary.

$$G_{\alpha,\beta}(i, j) = g_{\alpha}(i)g_{\beta}(j) \tag{5}$$

일반적으로 식 (5)와 같이 분리성(separability)을 만드

시 갖도록 할 필요는 없으나, 그 계산상 장점 때문에 일반적으로 식 (5)와 같은 분리된(separable) 사전을 사용한다. 식 (5)를 이용하여 만든 2차원 Gabor 사전이 그림 1에 나타나 있다.

2.2 Matching Pursuit

사전이 정해지면, 주어진 신호를 사전의 기본 신호들을 이용하여 식 (1)과 같이 표현하여야 한다. 이를 위하여 사용할 기본 함수와 각각에 대한 적절한 확장계수 값을 구하여야 하는데, 이를 위하여, 단계적으로 하나씩 최적의 기본 신호들을 찾아가는 greedy 알고리즘을 사용한다. 즉, 첫번째 분해 과정에서 식 (6)과 같이 원 신호를 가장 유사하게 표현할 수 있는 기본 신호 g_0 와 확장계수 a_0 를 구한다.

$$f(t) = a_0g_0(t) + R(t) \tag{6}$$

여기서 $R(t)$ 는 주어진 신호를 기본 신호 g_0 를 이용해 분해하고 남은 나머지 신호이며, a_0 는 임의의 기본 신호 g_0 가 주어졌을 때, 최소자승오차(minimum mean squared error)관점에서 주어진 신호를 최적 표현하는 계수값이다. 주어진 신호를 가장 잘 표현할 사전내의 한 기본 신호는 그 신호를 사용하여 원 신호를 근사화했을 때, 근사오차의 에너지가 가장 적어지는 함수이다. 따라서, 이 기본 신호함수와 해당 확장 계수값을 찾기 위하여 먼저 주어진 원 신호와 이의 근사 표현치간의 차신호 에너지를 표현하는 식 (7)의 내적을 정의한다.

$$E(a_0 | g_0) = \langle f - a_0g_0, f - a_0g_0 \rangle \tag{7}$$

식 (7)을 최소화하는 계수 a_0 를 찾으면 이는 최소자승 오차 관점에서 최적인 값이 되며, 이때 최소화된 $E(a_0 | g_0)$ 의 값은 식 (8)과 같다.

$$E(a_0 | g_0)_{min} = \langle f, f \rangle - \langle f, g_0 \rangle^2 \tag{8}$$

식 (7)을 최소화 하기 위하여 주어진 사전의 기본 신호들 중, $\langle f, g_0 \rangle^2$ 가 최대가 되는 기본 신호, 즉 입력신호와의 내적 절대값이 가장 큰 기본신호를 그 단계에서의 입력신호 분해를 위한 기본신호 g_0 로 한다. 이때의 확장계수 a_0 는 식 (7)을 최소화하기 위하여 다음과 같이 주어진다.

$$a_0 = \langle f, g_0 \rangle \quad \text{식 (9)}$$

식 (7) ~ (9)를 관찰하면, 이 과정은 결국 원 신호와 내적의 절대값이 가장 큰 기본 신호를 찾아내는 것으로, 개념적으로 원 신호와 가장 유사한 기본 신호들을 이용하여 주어진 입력 신호를 표현하는 것임을 알 수 있다. 주어진 입력신호를 분해한 후 근사오차 신호 $R(t)$ 를 이용하여 식 (10)과 같이 원 함수를 다시 정의한 후, 위 과정을 원하는 개수의 기본 신호로 원 신호를 분해할 때까지 순차적으로 반복하여 원 신호를 사전에 포함된 적절한 기본 신호들로 분해해 가는 것이 matching pursuit 알고리즘이다.

$$f(t) \leftarrow \{f(t) - a_0 g_0(t)\} \quad \text{식 (10)}$$

이와 같은 1차원 신호분해 과정은 2차원 영상에 대해서도 마찬가지로 적용됨을 쉽게 알 수 있다. 본 논문에서는 위의 matching pursuit과정을 2차원 움직임 보상 오차 영상에 적용하며, 이에 대한 순서도는 그림 2와 같다.

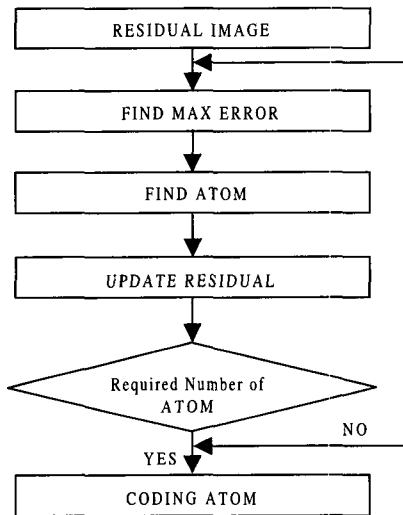


그림 2. Matching pursuit 과정
Fig. 2. Matching pursuit flowchart.

III. Matching Pursuit을 이용한 영상 부호화 방법과 기존의 고속 방법

3.1 Matching Pursuit을 이용한 영상 부호화 방법

그림 2에 나타난 것처럼 주어진 최초의 입력 영상인 움직임 보상 오차신호 영상을 가지고 matching pursuit 과정을 시작한다. 이때, 입력 영상인 움직임 보상 오차 신호 전체에 대해 2차원 기본 신호들간의 내적연산을 수행하는 대신 먼저, 12×12 크기의 슬라이딩 윈도우를 이용하여 에너지가 가장 많은 곳을 찾는 과정부터 수행한다^[4](그림 2의 “Find Max Error” 단계). 이는 움직임 보상 오차 영상의 많은 영역에서 에너지가 매우 작으며, 움직임 보상이 적절치 않은 곳에 주로 큰 에너지 신호가 있을 것이므로, 효율적인 부호화를 위해 전체 영상을 부호화하기 보다는 에너지가 높은 영역을 중점적으로 부호화 하는 것이 유리하기 때문이다. 또한 움직임 보상 오차신호의 특성상 부호화해야 할 부분이 실제로 듬성듬성 놓여져 있어 에너지가 가장 큰 부분부터 부호화하는 것이 효율적이기도 하다. 즉, 모든 영역에서 내적과정을 수행할 필요 없이 에너지가 가장 많이 몰려있는 부분을 찾고, 그 곳을 중심으로 2차원 신호를 분해하는 것이다. 이를 더욱 효율적으로 하기 위해, 12×12 크기의 슬라이딩 윈도우를 이용하여 최대 에너지가 있는 위치를 먼저 찾은 후, 이 위치를 기준으로 탐색영역 크기(S)를 16×16으로 확장한 후, 이 영역 내의 각 위치에 2차원 기본 신호를 대응시켜서 내적 연산을 수행한다^[4]. 이때 주어진 신호와 내적 연산의 절대치가 가장 큰 기본 신호가 선택되고, 그 신호의 인덱스와 영상내 해당 공간 좌표, 그리고 입력 영상과 선택된 신호의 내적 값 등 5가지 요소, 즉 선택된 atom정보를 전송 또는 저장한다. 그림 2의 순서도에 있어 “find atom”과정은 식 (7)~(9)의 내적연산을 위한 것이며, “Update Residual”단계는 선택된 기본 신호와 확장 계수를 곱해서 얻어지는 결과를 입력에서 제거하여 다음 단계에서 적용될 새로운 입력영상을 만드는 것이다.

3.2 기존의 고속 방법

Neff^[4]가 제안한 기존의 고속 방법은, Gabor 사전이 1차원의 기본 신호를 식 (5)를 이용하여 2차원으로 확장하여 만들어졌음을 이용하여 내적 연산을 수직과 수평 성분의 1차원 기본 신호로 나누어 계산한다. 그림 3은 2차원 입력 영상과 수직 성분의 1차원 기본 신호와의 내적 연산과, 그 결과로 얻어지는 중간 Matrix를 만드는 과정을 자세히 보여주고 있다. 그림 3에서처럼 2차원의 입력 영상과 수직 성분의 1차원 기본 신호의 내적 값을 구하면 그 결과로서 16×50의 중간 Matrix가 만들어 진다. 이렇게 만들어진 Matrix에 수평 성분

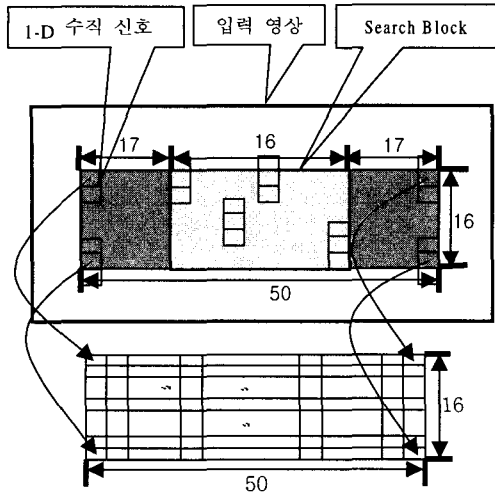


그림 3. 2차원 입력 영상과 1차원 기본 신호의 내적 과정
 Fig. 3. Inner product of 2-D input image and 1-D Gabor function.

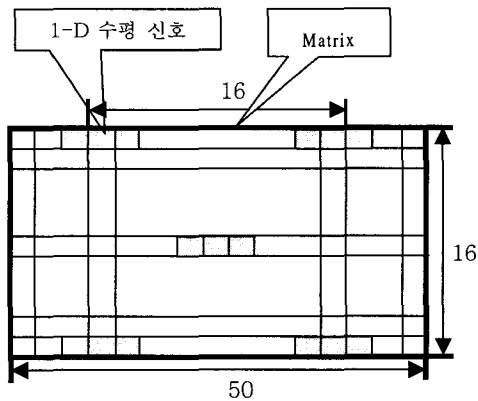


그림 4. Matrix와 가로방향의 1-D 기본 신호의 내적 연산
 Fig. 4. Inner product of matrix and 1-D Gabor function.

의 1차원 기본 함수를 다시 한번 내적함으로써 2차원 입력 영상과 2차원의 기본 신호와 내적이 이루어진다. 중간 Matrix와 수평 성분의 1차원 기본 신호와의 내적 연산과정은 그림 4에 나타나 있다.

IV. 제안하는 방법

본 논문에서 제안하는 방법은 세가지 특성을 이용하고 있는데, 그 첫번째는 1차원 사전을 구성하고 있는 각각의 기본 신호들의 절대값들이 대칭성을 가진다는

것이고, 두 번째는 크게 화질에 영향을 주지는 않지만 연산량을 줄이는 데는 도움을 줄 수 있도록 변형된 사전을 사용할 수 있다는 것이다. 이 두 가지 성질은 기본적으로 계산 방법을 달리하거나 매칭의 효율성을 높이기 위한 것이다. 마지막으로 이용하게 될 성질은 2차원 사전이 서로 다른 주파수 특성을 가진 부분들로 미리 나뉘어 질 수 있으므로, 주어진 입력신호를 분석하여 미리 정하여진 그룹의 어느 쪽에 속할 것인가를 결정하여 내적연산 대상의 기본 신호들의 수를 줄일 수 있다는 것이다. 이 방법은 앞서 기술한 두 가지 성질과 함께 사용할 수 있다.

4.1. 방법 1 : 절대값 대칭성을 이용

1차원 Gabor 사전의 기본 신호들을 살펴보면 그림 5와 같은 형태를 취하고 있음을 발견할 수 있다. 이것은 2장에서 알아본 바와 같이 1차원 Gabor 사전이 전형적인 우함수인 가우시안 윈도우 함수에 각각 스케일링과 평행이동, 변조를 가해 만들어졌기 때문이다. 따라서 이러한 형태의 값들에 대한 내적연산은 절대값들의 대칭성을 이용해서 더욱 간단히 수행할 수 있다. 본 논문에

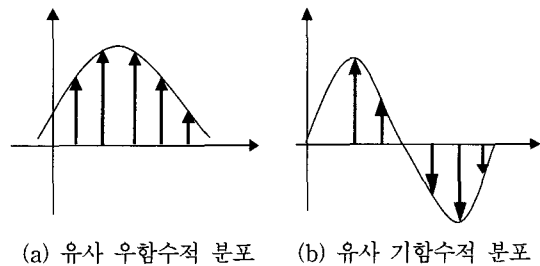


그림 5. 1차원 Gabor 사전의 형태
 Fig. 5. Shape of 1-D Gabor dictionary.

표 2. 유사 우함수의 예
 Table 2. Example of pseudo even function.

	0	...	m-1	m	...	2m-1	2m
$G[n]$	a	...	b	b	...	a	c
$F[n]$	e	...	f	g	...	h	i

표 3. 유사 기함수의 예
 Table 3. Example of pseudo odd function.

	0	...	m-1	m	...	2m-1	2m
$G[n]$	a	...	b	-b	...	-a	-c
$F[n]$	e	...	f	g	...	h	i

서는 그림 5의 두 가지 형태 함수의 중심이 원점이 아님을 구별하여 각각 유사 우함수, 유사 기함수라고 칭한다.

먼저, 유사 우함수의 경우에 대하여 연산량을 줄이는 과정을 살펴보자. 유사 우함수를 $G[n]$, 원 함수를 $F[n]$, 전체 샘플의 개수를 $2m+1$ 개라고 하고, $G[n]$ 과 $F[n]$ 이 표 2와 같이 값을 갖는다고 하자. 유사 우함수의 특성상, 전체 $2m+1$ 개의 함수 값들 중 마지막 하나의 함수 값 ($2m$ 번째 값)을 제외한 앞의 $2m$ 개의 함수 값들은 중앙을 기준으로 서로 같으므로 주어진 내적 연산을 식 (11)과 같이 계산할 수 있다.

$$\langle F, G \rangle = a(e+h) + \dots + b(f+g) + ci \quad \text{식 (11)}$$

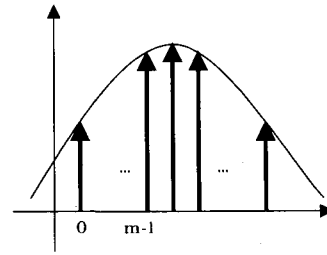
즉, 내적 연산 이전에 움직임 보상 오차신호 영상의 값들을 미리 더한 후, 내적 연산을 취하여 전체 내적 연산량을 줄일 수 있다. 마찬가지로 유사 기함수의 형태를 가질 때도, 표 3의 각 값들에 대해 식 (12)와 같이 거의 반의 곱셈 수를 사용하여 계산할 수 있다.

$$\langle F, G \rangle = a(e-h) + \dots + b(f-g) + (-c) \cdot i \quad \text{식 (12)}$$

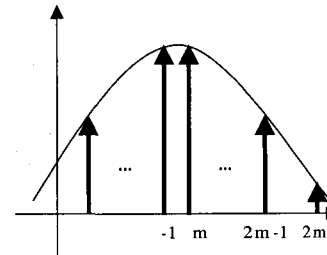
본 논문에서 사용할 Gabor 사전은 총 20개의 1차원 기본 신호들을 이용한 총 400개의 2차원 기본 신호 함수들로 구성되어 있는데, 20개의 1차원 기본 신호 함수들을 분석하면 표 1과 같이 13개의 유사 우함수와 5개의 유사 기함수, 그리고 대칭성이 전혀 없는 함수 2개로 이루어져 있음을 알 수 있다. 위의 대칭성을 이용한 방법은 각각의 1차원 기본함수에 적용되므로 종래의 분리성을 이용한 방법에 아무런 제약 없이 추가로 적용 가능하다.

4.2. 방법 2: 변형된 사전 이용

연속시간(continuous-time) 신호인 식 (3)의 Gabor 함수를 이산시간(discrete-time) Gabor 기본신호로 만들기 위해서는 표본화 과정을 수행하여야 한다. 이때, 표본화 과정을 어떻게 수행 하느냐에 따라 기본신호를 구성하는 각각의 값들이 달라지게 된다. 기존의 방법에서는 식 (4)를 구성하기 위하여 식 (3)의 u 값에 $\frac{N}{2}-1$ 을 사용하였다. 주어진 기본신호의 홀수개의 표본 값들이 좌우대칭을 이루도록 하려면 그림 6에 있는 두 가능성이 있는데, 이는 u 값에 $\frac{N}{2}-1$ 또는 $\frac{N-1}{2}$ 을 사용함으로써 각각 얻을 수 있다.



(a) 중점을 포함한 표본화 방법



(b) 중점을 포함하지 않은 표본화 방법

그림 6. 서로 다른 표본화 결과
Fig. 6. Results of different sampling.

이 두 가지 다른 표본화 방법 중 어느 것이 더 유리 할지는 matching pursuit을 수행할 실제 영상과 어느 것이 더 유사한가에 의해 결정된다. Neff의 Gabor 사전은 식 (4)처럼 $u = \frac{N}{2}-1$ 을 사용하였다. 한편, 입력 되는 신호 $F[n]$, 즉 움직임보상 오차영상의 에너지가 큰 영역을 실제로 관찰하면, 그림 6(b)의 예시같이 동일한 영상 값이 중앙의 한 점 이상에 걸쳐 분포한 경우가 많음을 볼 수 있다. 따라서 제안 방법에서는 기존 Gabor 사전의 기본 신호함수 값들을 구할 때, $u = \frac{N-1}{2}$ 을 사용하여, 위에 설명한 대칭성을 이용한 계산량 감축과 더불어 기본함수가 실제의 움직임 보상 오차영상에 더욱 근접하도록 하여 matching pursuit의 효율을 높이도록 하였다. 이를 위해 기본 신호함수의 길이를 고려하여, 절대값이 같은 중앙의 값들의 수를 다음과 같은 비례로 늘려가도록 하여, 방법 1에서 동시에 계산하는 개수를 늘리도록 하였다.

- 기본 신호 함수의 길이 = 3, 5 → 2개의 최대값
- 기본 신호 함수의 길이 = 7, 9, 11, 13, 15 → 3개의 최대값
- 기본 신호 함수의 길이 = 21, 23 → 4개의 최대값
- 기본 신호 함수의 길이 = 29, 35 → 5개의 최대값

예를 들어, 기본 신호의 길이가 7인 경우, 각각의 값은 표 4와 같이 바뀌어 진다. 실험 결과 방법 2는 PSNR측면에서 매우 미미하지만 약간의 이득도 있을 뿐더러 계산량을 줄일 수 있는 방법이므로 본 논문의 하나의 방법으로 사용되었다.

표 4. 변형시킨 기본신호의 예
Table 4. Example of modified bases.

기본신호	G[1]	G[2]	G[3]	G[4]	G[5]	G[6]	G[7]
변형전	A	B	C	D	E	F	G
변형후	A	B	sign(C)K	sign(D)K	sign(E)K	F	G
$K = \sqrt{\frac{C^2 + D^2 + E^2}{3}}$							

4.3 방법 3: 2차원 사전의 분류에 의한 방법

Matching pursuit은 원리상 주어진 함수와 가장 유사한 기본 신호들을 순차적으로 찾아내는 과정이다. 그러므로, 주어진 입력신호를 사전에 있는 모든 기본 신호와 전부 내적 연산하는 대신, 주어진 입력신호의 성질을 분석하여, 미리 분류된 기본 신호 그룹 중 어느 부류에 속하는 것과 가장 유사할 지를 판별한 후, 이 그룹내의 기본 신호들에 대해서만 내적연산을 취하도록 할 수 있다. 이 방법을 사용하기 위해 2차원 Gabor 사전을 살펴보면 그림 7에 보이는 것처럼 가운데를 중심으로 서로 다른 특징을 보이고 있는 다음의 네 가지 경우로 분류가 가능함을 알 수 있다.

영역 A : 저주파 성분이 많은 기본신호의 영역
영역 B : 수평 주파수성분이 많은 기본신호의 영역

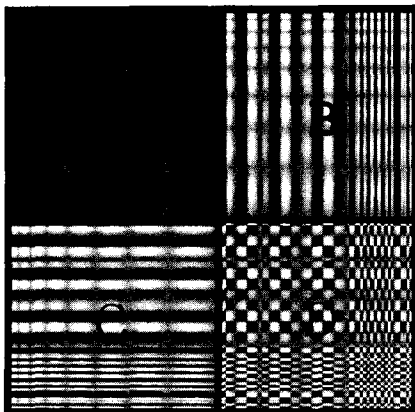


그림 7. 구분된 Gabor 사전
Fig. 7. Classified Gabor dictionary.

영역 C : 수직 주파수성분이 많은 기본신호의 영역
영역 D : 수직, 수평성분이 모두 많은 기본신호의 영역
이는 주어진 사전이 분리성을 갖고 있어 식 (5)처럼 수직 및 수평 주파수 성분으로 나뉠 수 있으며, 또한 표 1의 파라미터 값에 따라 주된 주파수 크기가 다르다는 것에 기인한다. 위에서 분류한 네 가지 영역으로 분류하는 것과 그 분류기준의 네 가지 영역에 대한 타당성을 확인하기 위하여 각 영역에 포함되어 있는 2차원 기본 신호들의 특징적인 주파수성분을 조사해 볼 필요가 있다. 이를 위해 본 논문에서는 되도록 최소한의 연산량으로 주파수성분에 따른 2차원 사전을 분류하기 위해, 식 (13)의 세가지 특징의 통계적 성질을 조사하였으며, 그 주파수 성분 분포가 그림 8에 있다.

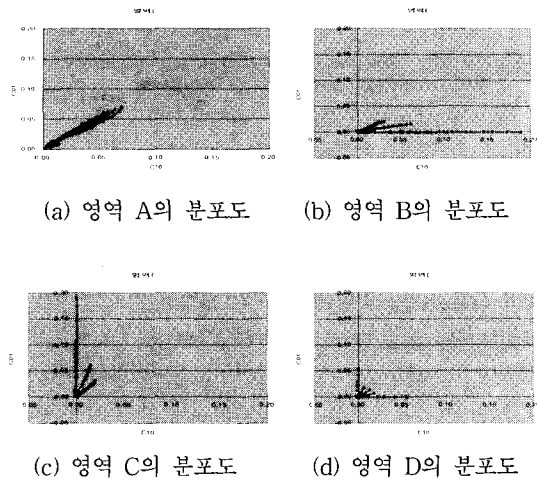


그림 8. 영역별 수직, 수평 주파수 분포도
Fig. 8. Distribution chart of vertical and horizontal frequencies.

식(13)은 주어진 SXS 탐색영역에 대해 한 번만 수행하게 된다. C10과 C01을 계산하기 위해 각각 (S+1)번의 곱셈이 필요하며 S=16을 사용하였으므로, 그 곱셈 수는 하나의 atom을 찾기 위해 소요되는 곱셈 수인 1,728,640번에 비교하여 (표 5 참조) 매우 작아 분류에 소요되는 오버헤드는 무시할 만하다.

$$C00 = \left| \sum_{x=0}^S \sum_{y=0}^S f(x, y) \right|$$

$$C10 = \frac{\sqrt{2}}{S} \left| \sum_{x=0}^S \sum_{y=0}^S f(x, y) \cos\left(\frac{(2x+1)\pi}{2S}\right) \right|$$

$$C01 = \frac{\sqrt{2}}{S} \left| \sum_{x=0}^{S-1} \sum_{y=0}^{S-1} f(x, y) \cos \frac{(2y+1)\pi}{2S} \right| \quad \text{식 (13)}$$

각 영역에 포함되어 있는 2차원 기본 신호들에 대한 대표적인 수직, 수평주파수 값의 절대치인 $C01$ 과 $C10$ 의 값들을 보면, 영역(A)의 경우 수직, 수평성분의 값이 통계적으로 유사하게 나타나고, 영역(B)의 경우는 수평주파수가 두드러지며, 영역(C)의 경우는 수직 주파수가 더 큰 것을 볼 수 있다. 영역(D)의 경우는 수직, 수평주파수가 고루 나타나지만 그 값이 크지 않음을 알 수 있다. 따라서 이러한 통계적 특성을 반영한 그림 9의 판단논리에 의해 2차원 사전에 포함되어 있는 모든 기본 신호를 네 영역으로 나누며, 이때 주어진 사전이 네 영역으로 구분되도록 파라미터 (α , β)값을 정한다.

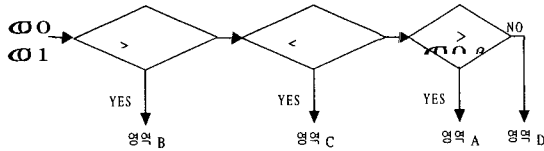


그림 9. 내적에 적용될 부분을 판단하는 과정
Fig. 9. Processing of selection in inner product.

그림 9는 주어진 신호의 $C00$, $C01$, $C10$ 값을 비교하여 해당 군을 찾는 분류순서이다. 먼저 $C01$ 과 $C10$ 의 값을 비교하여 $C10$ 이 $C01$ 의 α 배 보다 크면 네 영역 중에 (B)를 선택하고, 반대로 $C01$ 이 $C10$ 의 α 배 보다 작으면 (C)영역을 선택한다. 주어진 영상이 움직임 보상과정으로 인하여 발생한 오차 신호이며, 움직임 보상으로 발생하는 오차신호의 경우 일반적으로 양수와 음수의 값이 모두 존재할 수 있다. 따라서 (D)의 경우에서처럼 복잡한 형태의 영상은 양수와 음수가 서로 더해져서 사라지는 경향이 많아, 탐색영역내 모든 화소 값들의 $C00$ 의 값이 커질 수가 없으므로, $C01$ 과 $C10$ 의 값이 비슷할 경우에는 $C00$ 을 이용하여 (A)와 (D)를 구분할 수 있다. 즉, $C00$ 의 값이 어떤 임계 값 β 보다 크면 (A)를 선택하고, 작으면 (D)를 선택할 수 있다.

주어진 사전은 미리 네 영역으로 나뉘어져 있으므로, 실제 부호화 시 내적 연산을 수행 할 때는, matching pursuit을 수행할 입력신호에 대해서만 그림 9의 과정을 동일하게 적용시켜, 네 개의 군중 어느 하나로 결정한 후, 이에 해당하는 기본 신호들에 대해서만 내적연산을 하도록 하여 계산량을 1/4로 줄이도록 한다. 본

방법에 있어 선택된 군(subset)내의 내적연산 수행에도 역시 별도의 제약조건 없이 기존의 분리성 기반 방법^[4] 및 본 논문의 제안 방법 1과 방법 2를 같이 사용할 수 있다.

V. 실험결과

본 논문의 실험결과에 대한 구성은 먼저 실험에 사용된 조건을 밝히고, 위에서 제안한 세가지 방법 각각의 효율성을 먼저 비교한 후, 마지막으로 모든 과정에 대한 결과를 종합 나열하는 것이다. 방법(1)은 절대값의 대칭성을 이용한 방법이고, 방법(2)는 변형된 사전을 이용한 방법이며, 방법(3)은 주파수의 특성에 따라 분류된 사전 이용 방법이다. 비교 대상인 기존 방법은 Neff가 제안한 분리성을 이용한 고속 방법이며^[4], 계산량 비교는 하나의 atom을 찾기 위해 내적 연산에 사용되는 곱셈의 필요 수를 사용하였다.

실험조건

INTRA Coding에 사용된 QP = 20 (H.263 baseline)

10Kbps전송률을 가지는 채널을 가정하여 프레임 당 atom의 개수는 50개를 사용

Search Size(S) = 16

Threshold Value($\alpha = 3$, $\beta = 30$) 사용 (그림 9)

5.1. 방법 1 : 절대값 대칭성을 이용한 방법

본 논문에서 제안한 방법과 기존의 분리성만을 이용한 방법을 비교하기 위하여 먼저 분리성을 이용한 방법의 내적 연산량을 구해보면 식 (14)와 같다.

$$M = \sum_{v=1}^{20} \left\{ L_v \cdot S(S + L_{\max} - 1) + S^2 \sum_{h=1}^{20} L_h \right\} \quad \text{식 (14)}$$

여기서 L_v 는 v 번째 1-D Gabor 사전을 구성하는 기본 신호함수의 수직 크기를, 그리고 L_h 는 h 번째 1-D Gabor 사전의 기본 신호함수의 수평 크기를 의미한다. L_{\max} 는 L_v 또는 L_h 중 최대값을 의미하며, 실험에 사용한 Gabor 사전은 표 3과 같이 20개의 1차원 기본 함수를 기반으로 만들어졌다. 제안한 방법인 분리성과 대칭성을 이용한 방법을 함께 이용했을 때 연산량은 식 (15)와 같다.

$$M = \sum_{v=1}^{18} \left\{ \frac{L_v+1}{2} \cdot S \cdot L_{\max} - + S^2 \sum_{h=1}^{18} \frac{L_h+1}{2} + S^2 \sum_{h=19}^{20} L_h \right\} + \sum_{v=19}^{20} \left\{ L_v \cdot S \cdot S + L_{\max} - + S^2 \sum_{h=1}^{18} \frac{L_h+1}{2} + S^2 \sum_{h=19}^{20} L_h \right\}$$

식 (15)

실험 조건에 제시된 20개의 1차원 Gabor 사전을 이용하여 1개의 atom을 구하는데 필요한 연산량의 수치적인 비교는 표 5에 나타나 있다. 식 (15)와 표 5에서 보는 것처럼 영상의 화질에는 전혀 영향을 주지 않고, 내적에 필요한 곱셈 연산량을 거의 반으로 줄였음을 알 수 있다.

표 5. 방법(1)의 연산량 비교
Table 5. Comparison of total calculation using method (1).

	곱셈의 계산량	비교
기존의 방법	1,728,640	1.000
방법(1)	959,040	0.554

5.2. 방법 2: 변형된 사전 이용 방법

변형된 사전을 사용하는 두번째 방법은 첫번째 방법에 추가하여 계산량을 조금 더 줄이기 위한 것이다. 제안한 첫번째 방법과 두 번째 방법을 모두 적용했을 때의 연산량은 표 6과 같다. 그리고 이 방법을 적용했을 때의 화질의 PSNR변화는 표 7과 그림 10에 나타나 있다. 실제 부호화된 영상의 화질을 비교한 결과, 변형전

표 6. 방법(1)+(2)의 연산량 비교
Table 6. Comparison of total calculation using method (1)+(2).

	곱셈의 계산량	비교
기존의 방법	1,728,640	1.000
방법(1)+(2)	882,080	0.510

표 7. 변형된 사전의 PSNR 비교
Table 7. PSNR comparison of modified dictionary.

프레임	1	2	3	4	5	6	7	8	9	10
기존 방법	31.83	32.48	32.92	33.16	33.38	33.53	33.65	33.80	34.06	34.44
방법(1)+(2)	31.83	32.49	32.94	33.19	33.40	33.54	33.69	33.82	34.08	34.47

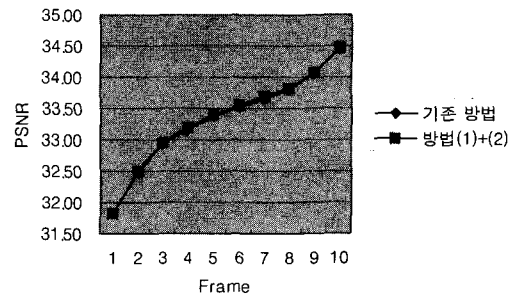


그림 10. PSNR 비교 [Akiyo, 프레임 1-10]
Fig. 10. Comparison of PSNR [Akiyo, frame 1-10].

의 사전을 사용했을 경우와 비교하여 PSNR면에서 약간의 개선이 있었으나, 육안으로는 거의 구별이 안 되는 화질 차이였다.

5.3 방법 3: 2차원 사전 분류에 의한 방법

주파수 특성에 의한 사전 분류 방법을 사용한다면 평균적으로 4개의 영역 중에 하나의 영역만을 내적연산의 대상으로 함으로 기존의 방법에 비해 계산량이 약 1/4로 감소한다. 그러나 분류과정이 완벽하지 못할 경우 약간의 PSNR의 감소를 수반할 수 있다. 그러나 계산측면에서 매우 간단한 그림 9의 분류방법을 사용하여도 그림 11의 Akiyo와 Mother & Daughter 영상에 대한 PSNR의 변화 그래프에서 나타나듯이 큰 화질 열화는 없다.

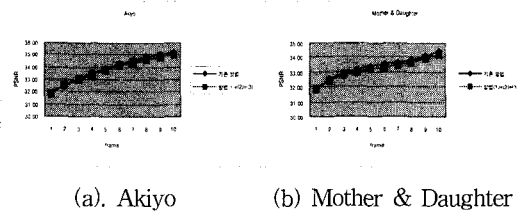


그림 11. PSNR 비교
Fig. 11. Comparison of PSNR.

5.4. 종합적인 결과

위에서 제안한 세 가지의 방법을 각각 (1)방법, (2)방법, (3)방법이라고 하고 모두 사용한 방법을 (1)+(2)+(3)방법이라 할 때, 전체적인 연산량과 평균 PSNR의 비교 결과는 표 8과 같다. 여기에서 비분리성(Non-Separable) 방법은 아무런 고속 연산 없이 2-D 내적연산을 그대로 수행하는 기존 고속 방법이다^[4].

표 8. 전체적인 방법들의 연산량과 PSNR비교
Table 8. Comparisons of total calculation and PSNR.

	연산량	연산량 비율(%)	평균 PSNR [dB]
Non-separable	21,827,584	100.00	33.28
Separable [4]	1,728,640	7.91	33.28
방법(1)	959,040	4.39	33.28
방법(1)+(2)	882,080	4.04	33.29
방법(1)+(2)+(3)	220,520	1.01	33.18

VI. 결 론

본 논문에서는 matching pursuit의 단점인 과도한 연산량의 문제점을 해결하고자 세가지 방법을 제시하였다. 첫번째 방법은 화질에는 전혀 열화를 주지 않고 연산량을 줄일 수 있으므로, 하드웨어 구현 시 적용된다면 효과를 얻을 것이다. 두 번째 방법은 화질 열화를 일으키지 않고, 연산량을 조금 더 줄일 수 있는 방법이다. 이 방법은 실험결과에서도 나타났듯이 영상에 적용했을 때 약간의 PSNR의 향상은 있었으나 무시할 만한 것이어서 이 보다는 연산량의 감소에 더 의미가 있는 방법이다. 세 번째 방법은 연산량을 1/4로 줄일 수 있으나 정확치 못한 분류에 따라 약간의 화질 열화가 수반될 가능성이 있다. 따라서 구하고자 하는 기본 신호 함수와 영상의 PSNR을 잘 고려하여 미미하지만 약간의 화질 열화를 감수하더라도 더 많은 수의 기본 신호 함수를 짧은 시간 내에 찾아내어 비디오 부호화에 적용하는 목적이려면 소기의 효과를 얻을 수 있을 것이다. 또한 본 논문에서 제안한 세가지 방법 모두 기존의 분리성을 이용한 고속 matching pursuit에 아무런 제약 조건 없이 부가하여 사용할 수 있는 장점이 있다. 한편, 본 논문은 Gabor 사전을 기반으로 한 고속 matching pursuit에 관한 것인데, 최근에 Gabor 사전 대신 subband 나 웨이블릿을 이용한 새로운 사전을 이용하여 계산량을 줄이는 방법도 새로이 연구되고 있다^[8]. 따라서, 본 논문의 제안 방법을 Gabor사전이 아닌 다른 경우에도 적용할 수 있도록 확장하는 것도 본 연구의 차후의 연구 방향이다.

참 고 문 헌

- [1] S. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Processing*, vol. 41, pp. 3397~3415, Dec. 1993.
- [2] F. Bergeaud and S. Mallat, "Matching pursuit of images," *Proc. IEEE-SP Int. Symp. Time-Frequency and Time-Scale Analysis*, pp. 330~333, Oct. 1994.
- [3] R. Neff, and A. Zakhor, "Matching pursuit video coding at very low bit rates," *IEEE Data Compression Conf., Snowbird, UT*, pp. 411~420, Mar. 1995.
- [4] R. Neff and A. Zakhor, "Very low bit rate video coding based on matching pursuits," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 7, pp. 158~171, Feb. 1997.
- [5] Mohammad Gharavi-Alkansari and Thomas S. Huang, "A Fast orthogonal matching pursuit algorithm," *IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP'98*, pp.1389~1392, May 1998.
- [6] Osama K. Al-Shaykh, Eugene Miloslavsky, Toshio Nomura, Ralph Neff, and Avideh Zakhor, "Video Compression Using Matching Pursuits," *IEEE Trans. Circuits Syst. Video Technol*, vol.9, pp. 123~143, Feb. 1999.
- [7] 오석병, 전병우, "벡터길이 비교를 이용한 고속 Matching Pursuit", *대한전자공학회 논문지*, 제38권 SP편 제2호, pp. 13~21, 2001
- [8] C. D. Vleeschouwer and B. Macq, "Subband dictionaries for low-cost matching pursuits of video residues," *IEEE Trans. Circuits Syst. Video Technol*, vol.9, no.7, pp. 984~993, Oct. 1999.

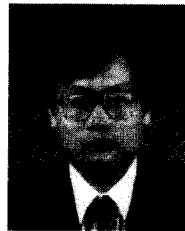
저 자 소 개



吳 錫 炳(正會員)

1999년 2월 : 성균관대학교 전자공학과 졸업(공학사). 2001년 2월 : 성균관대학교 전기전자 컴퓨터 공학부 졸업(공학석사). 현재 : LG전자 DM연구소 DVC그룹 IMT2000 팀 주임 연구원. <주관심분야 : 디

지털 비디오 처리, 멀티미디어>



孫 炳 宇(正會員)

1985년 2월 : 서울대학교 전자공학과 졸업(공학사). 1987년 : 서울대학교 전자 공학과 졸업(공학석사). 1992년 : Purdue Univ, School of Elec. 졸업(공학박사). 1993년~1997년 8월 : 삼성전자 멀티미디어

연구소 수석연구원. 1997년 9월~현재 : 성균관대학교 전기전자 컴퓨터 공학부 부교수. <주관심분야 : 멀티미디어, 영상압축, 영상인식, 영상처리>