



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2023-0032297
(43) 공개일자 2023년03월07일

(51) 국제특허분류(Int. Cl.)
G06T 3/40 (2006.01) G06T 3/60 (2006.01)
(52) CPC특허분류
G06T 3/4084 (2013.01)
G06T 3/60 (2013.01)
(21) 출원번호 10-2021-0114967
(22) 출원일자 2021년08월30일
심사청구일자 2021년08월30일

(71) 출원인
연세대학교 산학협력단
서울특별시 서대문구 연세로 50 (신촌동, 연세대학교)
(72) 발명자
최윤식
서울특별시 마포구 마포대로24길 16, 115동 204호 (아현동, 공덕자이 아파트)
이기환
서울특별시 서대문구 연세로 50, 제 3공학관 C60 3호(신촌동, 연세대학교)
(74) 대리인
특허법인(유한)아이시스

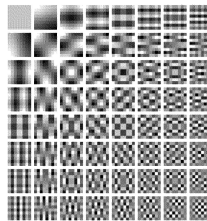
전체 청구항 수 : 총 10 항

(54) 발명의 명칭 기저 벡터 회전을 이용한 이산코사인변환 방법 및 영상 처리 장치

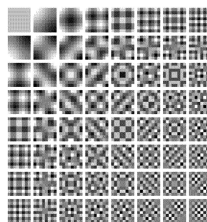
(57) 요약

기저 벡터 회전을 이용한 이산코사인변환 방법은 영상 처리 장치가 입력 영상을 복수의 패치들로 구분하는 단계, 상기 영상 처리 장치가 상기 복수의 패치들에 대응하는 DCT(discrete cosine transform) 매트릭스를 생성하는 단계, 상기 영상 처리 장치가 상기 패치들에서의 예지와 중심의 거리 및 DCT 계수의 부호를 기준으로 각도를 결정하는 단계, 상기 영상 처리 장치가 상기 각도에 따라 기저들을 회전할 회전 매트릭스를 생성하는 단계, 상기 영상 처리 장치가 상기 회전 매트릭스를 기준으로 변환 매트릭스를 생성하는 단계 및 상기 영상 처리 장치가 상기 변환 매트릭스를 이용하여 상기 입력 영상에 대한 DCT를 하는 단계를 포함한다.

대 표 도 - 도2



(A)



(B)

(52) CPC특허분류

H04N 19/625 (2015.01)

G06T 2207/20052 (2013.01)

명세서

청구범위

청구항 1

영상 처리 장치가 입력 영상을 복수의 패치들로 구분하는 단계;

상기 영상 처리 장치가 상기 복수의 패치들에 대응하는 DCT(discrete cosine transform) 매트릭스를 생성하는 단계;

상기 영상 처리 장치가 상기 패치들에서의 에지와 중심의 거리 및 DCT 계수의 부호를 기준으로 각도를 결정하는 단계;

상기 영상 처리 장치가 상기 각도에 따라 기저들을 회전할 회전 매트릭스를 생성하는 단계;

상기 영상 처리 장치가 상기 회전 매트릭스를 기준으로 변환 매트릭스를 생성하는 단계; 및

상기 영상 처리 장치가 상기 변환 매트릭스를 이용하여 상기 입력 영상에 대한 DCT를 하는 단계를 포함하는 기저 벡터 회전을 이용한 이산코사인변환 방법.

청구항 2

제1항에 있어서,

상기 영상 처리 장치가 상기 거리와 임계값을 비교한 결과 및 상기 DCT 계수 중 첫 번째 AC 계수들의 부호를 기준으로 상기 각도를 결정하는 기저 벡터 회전을 이용한 이산코사인변환 방법.

청구항 3

제1항에 있어서,

상기 영상 처리 장치가 아래 수식으로 상기 각도를 결정하는 기저 벡터 회전을 이용한 이산코사인변환 방법.

$$\theta = \begin{cases} \tan^{-1}\left(\left|\frac{C_{01}}{C_{10}}\right|\right) & \text{where } C_{01}C_{10} \geq 0 \text{ and } E_{low} > th \\ \tan^{-1}\left(\frac{\sqrt{\sum_{v=1}^n (C_{0v})^2}}{\sqrt{\sum_{u=1}^n (C_{u0})^2}}\right) & \text{where } C_{01}C_{10} \geq 0 \text{ and } E_{low} < th \\ 90^\circ - \tan^{-1}\left(\left|\frac{C_{01}}{C_{10}}\right|\right) & \text{where } C_{01}C_{10} < 0 \text{ and } E_{low} > th \\ 90^\circ - \tan^{-1}\left(\frac{\sqrt{\sum_{v=1}^n (C_{0v})^2}}{\sqrt{\sum_{u=1}^n (C_{u0})^2}}\right) & \text{where } C_{01}C_{10} < 0 \text{ and } E_{low} < th \end{cases}$$

(여기서, θ 는 상기 각도, $C_{01}C_{10}$ 은 상기 DCT 계수 중 첫 번째 AC 계수 2개의 곱, E_{low} 는 상기 거리이고, th 는 임계값)

청구항 4

제1항에 있어서,

상기 영상 처리 장치는 상기 기저들 중 일부 그룹에만 회전을 적용하는 기저 벡터 회전을 이용한 이산코사인변환 방법.

청구항 5

제4항에 있어서,

상기 일부 그룹은 상기 기저들 중 DC 계수를 제외한 첫 번째 열과 행인 제1 회전 그룹, 상기 기저들 대각 계수를 제외한 두 번째 열과 행인 제2 회전 그룹, 상기 제1 회전 그룹 및 상기 제2 회전 그룹을 포함하는 제3 회전 그룹 및 상기 기저들 DC 계수 및 대각 성분을 제외한 전체 기저들인 제4 회전 그룹 중 어느 하나의 그룹인 기저 벡터 회전을 이용한 이산코사인변환 방법.

청구항 6

기저 벡터 회전을 이용한 DCT(discrete cosine transform) 프로그램을 저장하는 저장 장치; 및

상기 프로그램을 이용하여 입력 영상을 압축하는 연산을 수행하는 연산장치를 포함하되,

상기 연산장치는 복수의 패치로 구분된 입력 영상에 대응하는 DCT 매트릭스를 생성하고, 상기 패치들에서의 위치와 중심의 거리 및 DCT 계수의 부호를 기준으로 각도를 결정하고, 상기 각도에 따라 기저들을 회전할 회전 매트릭스 및 변환 매트릭스를 생성하고, 상기 변환 매트릭스를 이용하여 상기 입력 영상에 대한 DCT를 수행하는 기저 벡터 회전을 이용한 이산코사인변환을 수행하는 영상 처리 장치,

청구항 7

제6항에 있어서,

상기 연산장치는 상기 거리와 임계값을 비교한 결과 및 상기 DCT 계수 중 첫 번째 AC 계수들의 부호를 기준으로 상기 각도를 결정하는 기저 벡터 회전을 이용한 이산코사인변환을 수행하는 영상 처리 장치,

청구항 8

제6항에 있어서,

상기 연산장치는 아래 수식으로 상기 각도를 결정하는 기저 벡터 회전을 이용한 이산코사인변환을 수행하는 영상 처리 장치,

$$\theta = \begin{cases} \tan^{-1}\left(\left|\frac{C_{01}}{C_{10}}\right|\right) \\ \text{where } C_{01}C_{10} \geq 0 \text{ and } E_{low} > th \\ \tan^{-1}\left(\frac{\sqrt{\sum_{v=1}^n (C_{0v})^2}}{\sqrt{\sum_{u=1}^n (C_{u0})^2}}\right) \\ \text{where } C_{01}C_{10} \geq 0 \text{ and } E_{low} < th \\ 90^\circ - \tan^{-1}\left(\left|\frac{C_{01}}{C_{10}}\right|\right) \\ \text{where } C_{01}C_{10} < 0 \text{ and } E_{low} > th \\ 90^\circ - \tan^{-1}\left(\frac{\sqrt{\sum_{v=1}^n (C_{0v})^2}}{\sqrt{\sum_{u=1}^n (C_{u0})^2}}\right) \\ \text{where } C_{01}C_{10} < 0 \text{ and } E_{low} < th \end{cases}$$

(여기서, θ 는 상기 각도, $C_{01}C_{10}$ 은 상기 DCT 계수 중 첫 번째 AC 계수 2개의 곱, E_{low} 는 상기 거리이고, th 는 임계값)

청구항 9

제6항에 있어서,

상기 연산장치는 상기 기저들 중 일부 그룹에만 회전을 적용하는 기저 벡터 회전을 이용한 이산코사인변환을 수행하는 영상 처리 장치,

청구항 10

제9항에 있어서,

상기 일부 그룹은 상기 기저들 중 DC 계수를 제외한 첫 번째 열과 행인 제1 회전 그룹, 상기 기저들 대각 계수

를 제외한 두 번째 열과 행인 제2 회전 그룹, 상기 제1 회전 그룹 및 상기 제2 회전 그룹을 포함하는 제3 회전 그룹 및 상기 기저들 DC 계수 및 대각 성분을 제외한 전체 기저들인 제4 회전 그룹 중 어느 하나의 그룹인 기저 벡터 회전을 이용한 이산코사인변환을 수행하는 영상 처리 장치,

발명의 설명

기술 분야

[0001] 이하 설명하는 기술은 기저 벡터 회전에 기반한 이산코사인변환 기법 및 영상 압축 기법이다.

배경 기술

[0002] 영상 압축 분야는 원본 정보의 손실을 최소화하면서 정보의 양을 줄이는 것이 주된 이슈이다. 이를 위하여 DCT(discrete cosine transform) 및 KLT(Karhunen-Loeve transform)와 같은 변환 기법이 활용된다. KLT는 가우시안 프로세스에 대한 최적의 변환 기법이지만 자원 소모가 많고 시간이 오래 소요된다.

[0003] DCT는 고속 처리가 가능하여 널리 사용되지만, 수평 또는 수직 방향의 에지 형태에 최적화되어있다. 따라서, DCT는 수평 또는 수직 방향이 아닌 다른 방향성에 대해서는 상대적으로 낮은 성능을 갖는다.

선행기술문헌

특허문헌

[0004] (특허문헌 0001) 미국공개특허 US2018-0338143호

발명의 내용

해결하려는 과제

[0005] 종래 DCT에서 기저(basis)를 회전하는 기법이 제안되었지만, 적합한 회전 정도를 추정하는 방법이 제시되지 못하였다.

[0006] 이하 설명하는 기술은 기저 회전을 위한 적절한 각도를 추정하여 DCT를 하는 기법을 제공하고자 한다.

과제의 해결 수단

[0007] 기저 벡터 회전을 이용한 이산코사인변환 방법은 영상 처리 장치가 입력 영상을 복수의 패치들로 구분하는 단계, 상기 영상 처리 장치가 상기 복수의 패치들에 대응하는 DCT 매트릭스를 생성하는 단계, 상기 영상 처리 장치가 상기 패치들에서의 에지와 중심의 거리 및 DCT 계수의 부호를 기준으로 각도를 결정하는 단계, 상기 영상 처리 장치가 상기 각도에 따라 기저들을 회전할 회전 매트릭스를 생성하는 단계, 상기 영상 처리 장치가 상기 회전 매트릭스를 기준으로 변환 매트릭스를 생성하는 단계 및 상기 영상 처리 장치가 상기 변환 매트릭스를 이용하여 상기 입력 영상에 대한 DCT를 하는 단계를 포함한다.

[0008] 기저 벡터 회전을 이용한 이산코사인변환을 수행하는 영상 처리 장치는 기저 벡터 회전을 이용한 DCT프로그램을 저장하는 저장 장치; 및 상기 프로그램을 이용하여 입력 영상을 압축하는 연산을 수행하는 연산장치를 포함한다. 상기 연산장치는 복수의 패치로 구분된 입력 영상에 대응하는 DCT 매트릭스를 생성하고, 상기 패치들에서의 에지와 중심의 거리 및 DCT 계수의 부호를 기준으로 각도를 결정하고, 상기 각도에 따라 기저들을 회전할 회전 매트릭스 및 변환 매트릭스를 생성하고, 상기 변환 매트릭스를 이용하여 상기 입력 영상에 대한 DCT를 수행한다.

발명의 효과

[0009] 이하 설명하는 기술은 낮은 복잡도로 기저의 회전 각도를 추정하여 최적의 DCT 변환을 가능하게 한다. 또한, 이하 설명하는 기술은 대부분의 영상 표준(JPEG, HEVC 등)에서 사용되는 DCT 변환 과정의 앞단에 단순한 각도 추정 구성만 추가하여 DCT 변환을 수행할 수 있다.

도면의 간단한 설명

- [0010] 도 1은 DCT의 기저들을 도시한 예이다.
- 도 2는 DCT의 기저 영상을 회전한 예이다.
- 도 3은 에지와 중심 사이의 거리가 다른 패치들을 도시한 예이다.
- 도 4는 기저 벡터를 복잡도 및 방향에 따라 분류한 예이다.
- 도 5는 4×4 블록 크기의 기저에 대하여 블록 그룹을 구분하여 회전한 결과이다.
- 도 6은 8×8 블록 크기의 기저에 대하여 블록 그룹을 구분하여 회전한 결과이다.
- 도 7은 4×4 크기의 DCT 매트릭스를 위한 회전 매트릭스의 예이다.
- 도 8은 영상 처리 장치에 대한 예이다.

발명을 실시하기 위한 구체적인 내용

- [0011] 이하 설명하는 기술은 다양한 변경을 가할 수 있고 여러 가지 실시례를 가질 수 있는 바, 특정 실시례들을 도면에 예시하고 상세하게 설명하고자 한다. 그러나, 이는 이하 설명하는 기술을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 이하 설명하는 기술의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0012] 제1, 제2, A, B 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 해당 구성요소들은 상기 용어들에 의해 한정되지는 않으며, 단지 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다. 예를 들어, 이하 설명하는 기술의 권리 범위를 벗어나지 않으면서 제1 구성요소는 제2 구성요소로 명명될 수 있고, 유사하게 제2 구성요소도 제1 구성요소로 명명될 수 있다. 및/또는 이라는 용어는 복수의 관련된 기재된 항목들의 조합 또는 복수의 관련된 기재된 항목들 중의 어느 항목을 포함한다.
- [0013] 본 명세서에서 사용되는 용어에서 단수의 표현은 문맥상 명백하게 다르게 해석되지 않는 한 복수의 표현을 포함하는 것으로 이해되어야 하고, "포함한다" 등의 용어는 설명된 특징, 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것이 존재함을 의미하는 것이지, 하나 또는 그 이상의 다른 특징들이나 개수, 단계, 동작, 구성요소, 부분품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 배제하지 않는 것으로 이해되어야 한다.
- [0014] 도면에 대한 상세한 설명을 하기에 앞서, 본 명세서에서의 구성부들에 대한 구분은 각 구성부가 담당하는 주기능 별로 구분한 것에 불과함을 명확히 하고자 한다. 즉, 이하에서 설명할 2개 이상의 구성부가 하나의 구성부로 합쳐지거나 또는 하나의 구성부가 보다 세분화된 기능별로 2개 이상으로 분화되어 구비될 수도 있다. 그리고 이하에서 설명할 구성부 각각은 자신이 담당하는 주기능 이외에도 다른 구성부가 담당하는 기능 중 일부 또는 전부의 기능을 추가적으로 수행할 수도 있으며, 구성부 각각이 담당하는 주기능 중 일부 기능이 다른 구성부에 의해 전담되어 수행될 수도 있음은 물론이다.
- [0015] 또, 방법 또는 동작 방법을 수행함에 있어서, 상기 방법을 이루는 각 과정들은 문맥상 명백하게 특정 순서를 기재하지 않은 이상 명기된 순서와 다르게 일어날 수 있다. 즉, 각 과정들은 명기된 순서와 동일하게 일어날 수도 있고 실질적으로 동시에 수행될 수도 있으며 반대의 순서대로 수행될 수도 있다.
- [0017] 이하 설명하는 기술은 최적의 DCT 기법에 관한 것이다. 이하 설명하는 부가적인 최적화 과정 없이 최적 DCT를 위한 기저 벡터의 회전 각도를 결정한다.
- [0019] 일반적으로 DCT는 압축 도메인에서 영상의 정보를 추출하고 분석하는 것이 중요하다.
- [0020] 도 1은 DCT의 기저들(bases)을 도시한 예이다. 도 1은 2차원 DCT(2D-DCT)의 기저들을 나타낸다. DCT의 기저들은 수평 방향 및 수직 방향 또는 수평과 수직의 조합으로 형성되는 대각 방향이다. 두 개의 방향성 기저들은 순서에 따라 동일한 에지 복잡도를 갖는다. 도 1에서 사각형 박스로 표시한 기저들은 서로 다른 방향에서 동일한 복잡도의 에지 정보를 갖는다. DCT는 도 1에서의 수직 정보와 수평 정보에 대하여 좋은 성능을 갖는다고 알려졌다.

[0021] 예지 방향성은 아래 수학식 1 내지 수학식 4로 표현될 수 있다. 아래 수식은 8×8 블록 DCT를 기준으로 계수를 설명한 것이다.

수학식 1

$$\tan \theta = \left(\sum_{v=1}^7 C_{0v} \right) / \left(\sum_{u=1}^7 C_{u0} \right)$$

[0023]

수학식 2

$$\tan \theta = \left(\sum_{v=1}^7 H_{0v} \right) / \left(\sum_{u=1}^7 H_{u0} \right)$$

[0024]

수학식 3

$$\tan \theta = C_{01} / C_{10}$$

[0026]

수학식 4

$$\tan \theta = H_{01} / H_{10}$$

[0028]

[0029] C_{0v} 및 C_{u0} 은 도 1에서 각각 첫 번째 행과 열의 기저들에 대한 DCT 계수를 나타낸다. H_{0v} 및 H_{u0} 은 허프만 디코딩을 통하여 압축된 스트림에서 추출한 계수를 나타낸다.

[0031] 도 2는 DCT의 기저 영상을 회전한 예이다. 도 2는 방향성 에지를 나타내는 특정 각도를 회전한 결과이다. 종래 연구 SDCT(steerable DCT, G. Fracastoro et al., "Steerable Discrete Cosine Transform," 2015 IEEE 17th International Workshop on Multimedia Signal Processing (MMSP), 2015, 등 참고) 기법은 방향성 에지를 나타내는 아래 수학식 5를 사용하여 정해진 각도로 기저를 회전시키었다.

수학식 5

$$\begin{bmatrix} \mathbf{v}^{(k,l)'} \\ \mathbf{v}^{(l,k)'} \end{bmatrix} = \begin{bmatrix} \cos \theta_{k,l} & \sin \theta_{k,l} \\ -\sin \theta_{k,l} & \cos \theta_{k,l} \end{bmatrix} \begin{bmatrix} \mathbf{v}^{(k,l)} \\ \mathbf{v}^{(l,k)} \end{bmatrix}$$

[0033]

[0034] 여기서 $0^\circ \leq \theta^{(k,1)'} \leq 90^\circ$ 이고, $\mathbf{v}^{(k,1)'}$ 및 $\mathbf{v}^{(1,k)'}$ 는 SDCT에서의 기저들을 의미한다.

[0035] 도 2는 상기 수학식 5를 사용하여 결정한 각도를 회전한 결과라고 할 수 있다. 도 2(A)는 기저 영상 각 쌍에 대

하여 동일하게 20° 를 회전한 예이고, 도 2(B)는 기저 영상 각 쌍에 대하여 동일하게 45° 를 회전한 예이다.

- [0037] 종래 DCT 기저를 회전하는 방식은 반복적인 과정을 통해 최적의 회전값을 찾는 방식이었다. 이 경우 DCT에 많은 시간이 소요될 수 있기에 실제 영상 애플리케이션에 적용하기 어려웠다.
- [0038] 2D-DCT가 충분한 방향 정보를 갖고 있다면, DCT를 위한 기저의 회전 값을 결정할 수 있다. 종래 연구(J. Hou et al., "Sparsifying orthogonal transforms with compact bases for data compression," 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, 2016)는 기저 매트릭스(basis matrix)와 회전 매트릭스(rotation matrix)를 곱하여 산출되는 변환 매트릭스(transform matrix)를 제시한 바 있다. 다만, n차원 회전 매트릭스 생성은 복잡한 최적화를 필요로 한다고 알려졌다. 따라서, 복잡도 낮은 회전 매트릭스 생성 방법이 필요하다.
- [0039] 연구자는 적은 개수의 각도들을 사용하여 에너지 압축(energy compaction) 관점에서 KLT와 유사하게 효과적인 변환 매트릭스를 생성하는 방법을 제안한다.
- [0041] 이하 영상 처리 장치가 DCT를 수행한다고 가정한다. 영상 처리 장치는 영상 코딩 장치하는 장치일 수 있다.
- [0042] 영상 처리 장치는 입력되는 영상에서 방향성을 추정하고 기저를 회전하여 DCT를 수행할 수 있다. 이하 구체적인 과정을 설명한다.
- [0043] 영상 처리 장치는 2D-DCT의 특징을 기준으로 주방향(main orientation)을 추정한다. 도 1에서 설명한 바와 같이 2D-DCT는 수직 및 수평 방향에 따라 일정한 방향성 정보를 포함한다. 방향성 정보는 수학적 1 내지 수학적 4를 이용하여 확인할 수 있다.
- [0044] 상기 수학적 1 내지 수학적 4에 기반하여 각도들을 추정하기 위한 과정을 설명한다. 영상 처리 장치는 아래 과정을 통해 기저 회전을 위한 각도를 연산할 수 있다.

수학적 6

$$E_{low} = \frac{\sqrt{\sum_{(u,v) \in [0,1]} C_{uv}^2}}{\sqrt{\sum_{(u,v) \in [0,n-1]} C_{uv}^2}}$$

- [0046]
- [0047] 수학적 6은 패치(patch)의 에지와 중심 사이의 거리 (E_{low})를 나타낸다.
- [0048] 도 3은 에지와 중심 사이의 거리가 다른 패치들을 도시한 예이다. 도 3(A) 내지 도 3(D)는 각각 E_{low} 가 0.8813, 0.8974, 0.9926 및 0.9917이다. 도 3(A)와 도 3(B)는 E_{low} 가 0.90보다 낮고, 에지들이 중심에서 멀리 위치한다. 이와 달리 도 3(C) 및 도 3(D)는 상대적으로 높은 E_{low} 를 가지며 에지가 패치의 중심 근처에 위치한다.
- [0049] 각 기저는 이미 90° 회전한 쌍을 갖기 때문에 각도 범위는 0° ~ 90° 로 제한된다. DCT 계수들의 부호(sign)는 대응하는 기저와 함께 중심으로부터 에지의 방향(왼쪽 또는 오른쪽)을 나타낸다. 따라서 첫 번째 AC 계수들의 부호를 이용하여 방향의 범위를 제한할 수 있다. 연구자는 실험적으로 E_{low} 의 임계값 th 을 0.90으로 설정하고, 아래 수학적 7과 같이 방향을 수식화하였다. 즉, 영상 처리 장치는 첫 번째 AC 계수 2개의 곱 $C_{01}C_{10}$ 과 E_{low} 과 임계값 th 를 비교한 결과에 따라 각도를 결정한다.

수학식 7

$$\theta = \begin{cases} \tan^{-1}\left(\left|\frac{C_{01}}{C_{10}}\right|\right) \\ \quad \text{where } C_{01}C_{10} \geq 0 \text{ and } E_{low} > th \\ \tan^{-1}\left(\frac{\sqrt{\sum_{v=1}^n (C_{0v})^2}}{\sqrt{\sum_{u=1}^n (C_{u0})^2}}\right) \\ \quad \text{where } C_{01}C_{10} \geq 0 \text{ and } E_{low} < th \\ 90^\circ - \tan^{-1}\left(\left|\frac{C_{01}}{C_{10}}\right|\right) \\ \quad \text{where } C_{01}C_{10} < 0 \text{ and } E_{low} > th \\ 90^\circ - \tan^{-1}\left(\frac{\sqrt{\sum_{v=1}^n (C_{0v})^2}}{\sqrt{\sum_{u=1}^n (C_{u0})^2}}\right) \\ \quad \text{where } C_{01}C_{10} < 0 \text{ and } E_{low} < th \end{cases}$$

[0051]

[0053] 두 개의 행 벡터들로 형성되는 공간 회전은 주어진 회전 매트릭스로 정의될 수 있다. 수학식 5는 2개의 벡터의 회전을 설명하며, 아래 수학식 8과 같이 n^2 차원 형태로 표현될 수 있다.

수학식 8

$$R_{ij} = \begin{bmatrix} I_{i-1} & & & & \\ & \cos \theta & & \sin \theta & \\ & & I_{j-i-1} & & \\ & -\sin \theta & & \cos \theta & \\ & & & & I_{n^2-j} \end{bmatrix} \in \mathbb{R}^{n^2 \times n^2}$$

[0055]

[0056] 여기서, i 및 j 는 행 벡터의 인덱스이고, n 는 블록의 크기이다.

[0058] 연구자는 각 패치를 특정 각도로 회전시킬 최적의 쌍이 있다는 것을 발견하였다. 연구자는 사전 연산된 각도에 따라 기저 벡터를 복잡도와 방향이 서로 다른 5개의 그룹으로 분류하였다. 도 4는 기저 벡터를 복잡도 및 방향에 따라 분류한 예이다. 도 4(A)는 8×8 크기를 갖는 경우이고, 도 4(B)는 4×4 크기를 갖는 경우이다. 도 4는 복잡도와 방향에 따라 서로 다른 그룹을 서로 다른 색으로 표현하였다. 도 4에서 짙은 파란색(navy)은 DC 계수를 나타내고, 옅은 파란색은 저주파수를 갖는 기저를 나타내고, 황색(orange)은 대각 성분을 나타내고, 노란색은 두 번째 레벨의 기저를 나타내고, 회색은 나머지 기저들을 표현하였다.

[0059] 연산된 각도는 주된 방향성을 나타내지만 개별 패치 내의 복잡한 상태를 포함하지는 않는다. 따라서 일정한 경우 저주파수 요소만을 회전시키는 것이 보다 높은 성능을 나타낼 수 있다. 연구자는 크기 8×8 및 4×4 에 대하여 최적의 기저 쌍을 도출하였다.

[0060] 최적의 기저 쌍은 식별하기 위하여, 연구자는 다음과 같은 4개의 그룹을 분류하였다. (1) 회전 그룹 1은 DC 계수를 제외한 첫 번째 열과 행(도 4의 옅은 파란색 블록)이고, (2) 회전 그룹 2는 대각 계수를 제외한 두 번째 열과 행(도 4의 노란색 블록)이고, (3) 회전 그룹 3은 회전 그룹 1과 회전 그룹 2(도 4에서 옅은 파란색 블록 및 노란색 블록)이고, (4) 회전 그룹 4는 DC 계수 및 대각 성분을 제외한 전체 기저 쌍(도 4에서 짙은 파란색, 노란색 및 회색 블록)이다.

[0061] 연구자는 테스트 영상을 이용하여 회전 그룹에 따라 선별적으로 기저를 회전하여 DCT한 결과를 확인하였다. 도 5는 4×4 블록 크기의 기저에 대하여 블록 그룹을 구분하여 회전한 결과이다. 도 6은 8×8 블록 크기의 기저에

대하여 블록 그룹을 구분하여 회전한 결과이다. 도 5 및 도 6에서 각각 (A), (B), (C) 및 (D)는 테스트 영상 Barbara, peppers, Lena 및 house를 각각 사용한 결과이다. 도 5 및 도 6은 보유한 계수의 개수와 PSNR(dB)의 관계를 도시한다. 도 5 및 도 6에서 검은색 선은 회전 그룹 1을 나타내고, 보라색 선은 회전 그룹 2를 나타내며, 파란색 선은 회전 그룹 3을 나타내고, 붉은색 선은 회전 그룹 4를 나타낸다.

[0062] 도 5를 살펴보면 회전 그룹 4인 경우 가장 성능이 높았고, 다음으로는 회전 그룹 3의 성능이 높았다. 도 6을 살펴보면 기저 개수가 적은 경우 회전 그룹 3이 가장 성능이 높았고, 기저 개수가 많아지면 회전 그룹 1의 성능이 높아지는 경향을 보인다. 따라서 일반적으로 회전 그룹 3이 최적의 그룹이라고 할 수 있다. 따라서, 회전 대상이 되는 특정 블록을 선별하여 DCT 하는 방식이 효과적이라는 것을 알 수 있다. 나아가 영상의 특징을 실험적으로 결정하고, 영상의 특징이나 종류에 따라 회전할 최적의 기저 쌍 세트를 선별할 수도 있을 것이다.

[0063] 정리하면, 영상 처리 장치는 사전에 설정되는 정보에 따라 최적의 기저 쌍 세트를 선별하여 회전 대상으로 삼을 수 있다. 이를 회전시킬 최적 기저 쌍 세트라고 하면, 아래 수학적 9는 전체 회전 매트릭스를 나타낸다. 여기서, i 및 j 는 벡터들의 인덱스이다.

수학적 9

$$R(\theta) = \prod_{(i,j) \in \Omega} R_{ij}$$

[0065]

[0067] 도 7은 4×4 크기의 DCT 매트릭스를 위한 회전 매트릭스의 예이다. 회전 매트릭스는 θ 값만 알려지면 부가적인 연산 없이 즉각적으로 생성될 수 있다. 이후 영상 처리 장치는 아래 수학적 10과 같이 회전 매트릭스와 2D-DCT 매트릭스를 곱하여 변환 매트릭스 H 를 쉽게 마련할 수 있다.

수학적 10

$$H = R(\theta)T$$

[0069]

[0070] T 는 2D-DCT 매트릭스이고, θ 는 수학적 7에서 결정되는 각도이다.

[0072] 전술한 최적의 기저 블록 쌍 세트를 선택적으로 회전시키는 DCT(partially rotated DCT)는 아래 표 1과 같이 정리될 수 있다.

표 1

Algorithm 1: Implementation of Partially Rotated Discrete Cosine Transform

Input: Input image X , Block size n

Output: Transform matrix H

1. Convert image X to $n \times n$ patches x ;
2. Set T to $n \times n$ DCT matrix;
3. Compute the DCT coefficients, $C = T x T^T$;
4. Compute E_{low} using ;
5. Compute the angle θ ;
6. Compute the rotation matrix $R(\theta)$;
7. Apply the rotation matrix $R(\theta)$;

[0074]

[0075] 영상 처리 장치는 압축 대상인 입력 영상 X 를 $n \times n$ 크기의 패치로 구분(분할)한다. 영상 처리 장치는 $n \times n$ 크기의 DCT 매트릭스를 설정하고, DCT 계수를 연산한다. DCT 매트릭스를 생성하고, 이후 영상 처리 장치는 변환 매트릭스를 생성한다.

[0076] 영상 처리 장치는 변환 매트릭스 생성을 위하여 회전 매트릭스를 생성한다. 영상 처리 장치는 수학식 6과 같이 패치에서 중심과 에지 사이의 거리 E_{low} 를 계산한다. 영상 처리 장치는 DCT 계수 및 E_{low} 를 기준으로 수학식 7과 같이 회전 각도 θ 를 결정한다. 영상 처리 장치는 θ 에 따라 회전 매트릭스를 생성하고, 회전 매트릭스와 DCT 매트릭스를 곱하여 변환 매트릭스를 생성한다. 영상 처리 장치는 입력 영상의 패치에 변환 매트릭스를 적용하여 DCT를 수행한다.

[0078] 아래 표 2는 제안한 DCT와 종래 기법의 복잡도를 비교한 것이다. 표 2는 복잡도는 $n \times n$ 크기 패치를 갖는 영상을 대상으로 하고, T 는 SOT에서의 수렴위한 반복 횟수, s 는 목표하는 희소도(sparsity)이고, Q 는 양자화 레벨을 나타낸다.

표 2

Method	Number of operations	Order of complexity
DCT	$2n(n \log n - n + 2)N$	$O(n^2 N \log n)$
KLT	$3n^4 N + 9n^6 + 2n^2 N$	$O(Nn^4)$ or $O(n^6)$
SOT	$(2n^4 N + n^2 s N + 21n^6)T + (2n^4 - n^2)N$	$O(n^4 NT)$ or $O(n^6 T)$
SDCT	$(2n(n \log n - n + 2) + n^3)QN$	$O(n^3 QN)$
Ours	$(4n^2 \log n + n^3 + 8n + 1)N$	$O(n^3 N)$

[0079]

[0080] 표 2는 전통적인 DCT, KLT, SOT(sparse orthonormal transform), SDCT 및 전술한 제안 기술(Ours)을 비교한 결과이다. 표 2에서 DCT는 1차원 DCT를 의미한다. 표 2를 살펴보면, 1차원 DCT를 제외하면 다른 기법에 비하여 제안한 방법의 복잡도가 낮은 것을 알 수 있다. 종래 영상 압축 성능을 높이기 위한 기저 회전 방식은 복잡도가 높아 실제 영상 애플리케이션에 적용하기 어려웠지만, 제안 기술은 간단하게 회전 각도를 결정하여 복잡도 낮게 DCT가 가능하다.

[0082] 도 8은 영상 처리 장치(100)에 대한 예이다. 영상 처리 장치는 영상을 압축하는 인코딩 장치일 수 있다.

[0083] 영상 처리 장치(100)는 영상 압축을 위한 DCT를 하는 장치이다. 영상 처리 장치(100)는 물리적으로 다양한 형태로 구현될 수 있다. 예컨대, 영상 처리 장치(100)는 스마트 기기, 컴퓨터 장치, 네트워크의 서버, 영상 처리 전용 칩셋 등의 형태를 가질 수 있다.

[0084] 영상 처리 장치(100)는 저장 장치(110), 메모리(120), 연산장치(130), 인터페이스 장치(140) 및 통신 장치(150)

0)를 포함할 수 있다.

- [0085] 저장 장치(110)는 영상 압축을 위한 프로그램 내지 코드를 저장한다. 저장 장치(110)는 전술한 DCT를 위한 프로그램 내지 코드를 저장할 수 있다.
- [0086] 저장 장치(110)는 기타 데이터 처리에 필요한 프로그램 내지 소스 코드 등을 저장할 수 있다. 저장 장치(110)는 압축 대상이 영상 데이터 및 압축된 영상 데이터 등을 저장할 수 있다.
- [0087] 메모리(120)는 DCT 및 영상 압축 과정에서 생성되는 데이터 및 정보 등을 저장할 수 있다. 예컨대, 메모리(120)는 전술한 E_{low} , 회전 각도, 회전 대상인 기저 쌍 세트 등을 저장할 수 있다.
- [0088] 인터페이스 장치(140)는 외부로부터 일정한 명령 및 데이터를 입력받는 장치이다. 인터페이스 장치(140)는 물리적으로 연결된 입력 장치 또는 외부 저장 장치로부터 압축 대상인 영상을 입력받을 수 있다. 인터페이스 장치(140)는 전술한 DCT를 위하여 사전에 필요한 정보(수학식 7에서 각도 결정을 위한 거리 임계값, 회전 대상인 기저 쌍 세트 등)를 입력받을 수도 있다. 인터페이스 장치(140)는 압축한 영상을 외부 객체에 전송할 수도 있다.
- [0089] 통신 장치(150)는 유선 또는 무선 네트워크를 통해 일정한 정보를 수신하고 전송하는 구성을 의미한다. 통신 장치(150)는 외부 객체로부터 압축 대상인 영상을 수신할 수 있다. 통신 장치(150)는 전술한 DCT를 위하여 사전에 필요한 정보(수학식 7에서 각도 결정을 위한 거리 임계값, 회전 대상인 기저 쌍 세트 등)를 수신할 수도 있다. 통신 장치(150)는 압축한 영상을 외부 객체에 송신할 수 있다.
- [0090] 연산 장치(130)는 저장장치(110)에 저장된 프로그램 내지 코드를 이용하여 DCT 및 영상 압축을 한다.
- [0091] 연산 장치(130)는 입력 영상을 $n \times n$ 크기의 패치로 구분한다.
- [0092] 연산 장치(130)는 $n \times n$ 크기의 패치에 대응하는 DCT 매트릭스를 생성한다. 이 과정에서 연산 장치(130)는 DCT 계수를 결정한다. DCT 계수 연산이나 DCT 매트릭스 생성은 영상 압축 분야에서 사용되는 다양한 방법 중 어느 하나가 사용될 수 있다.
- [0093] 연산 장치(130)는 DCT를 위한 변환 매트릭스를 생성한다. 이를 위하여 연산 장치(130)는 회전 매트릭스를 먼저 생성한다.
- [0094] 연산 장치(130)는 일정한 기저들을 회전하기 위하여, 수학식 6을 이용하여 패치에서 중심과 에지의 거리를 연산한다. 연산 장치(130)는 수학식 7과 같이 연산한 거리와 거리 임계값을 비교한 결과 및 DCT 계수를 기준으로 각도 θ 를 결정한다. 거리 임계값은 애플리케이션, 영상 종류 등에 따라 사전에 결정된 값일 수 있다.
- [0095] 연산 장치(130)는 결정된 각도 θ 를 기준으로 회전 매트릭스를 생성한다. 도 7에 도시한 바와 같이 θ 만 결정되면 회전 매트릭스는 즉시 생성 가능하다.
- [0096] 연산 장치(130)는 회전 매트릭스와 2D-DCT 매트릭스를 곱하여 변환 매트릭스를 생성한다.
- [0097] 연산 장치(130)는 생성된 변환 매트릭스를 이용하여 DCT를 수행한다.
- [0098] 나아가, 연산 장치(130)는 DCT 이후 영상 압축에 필요한 다른 과정(양자화 등)을 수행할 수도 있다.
- [0099] 연산 장치(130)는 데이터를 처리하고, 일정한 연산을 처리하는 프로세서, AP, 프로그램이 임베디드된 칩과 같은 구성일 수 있다.
- [0101] 또한, 상술한 바와 같은 DCT 방법은 컴퓨터에서 실행될 수 있는 실행가능한 알고리즘을 포함하는 프로그램(또는 어플리케이션)으로 구현될 수 있다. 상기 프로그램은 일시적 또는 비일시적 판독 가능 매체(non-transitory computer readable medium)에 저장되어 제공될 수 있다.
- [0102] 비일시적 판독 가능 매체란 레지스터, 캐쉬, 메모리 등과 같이 짧은 순간 동안 데이터를 저장하는 매체가 아니라 반영구적으로 데이터를 저장하며, 기기에 의해 판독(reading)이 가능한 매체를 의미한다. 구체적으로는, 상술한 다양한 어플리케이션 또는 프로그램들은 CD, DVD, 하드 디스크, 블루레이 디스크, USB, 메모리카드, ROM(read-only memory), PROM(programmable read only memory), EPROM(Erasable PROM, EPROM) 또는 EEPROM(Electrically EPROM) 또는 플래시 메모리 등과 같은 비일시적 판독 가능 매체에 저장되어 제공될 수 있다.
- [0103] 일시적 판독 가능 매체는 스태틱 램(Static RAM, SRAM), 다이내믹 램(Dynamic RAM, DRAM), 싱크로너스 디램

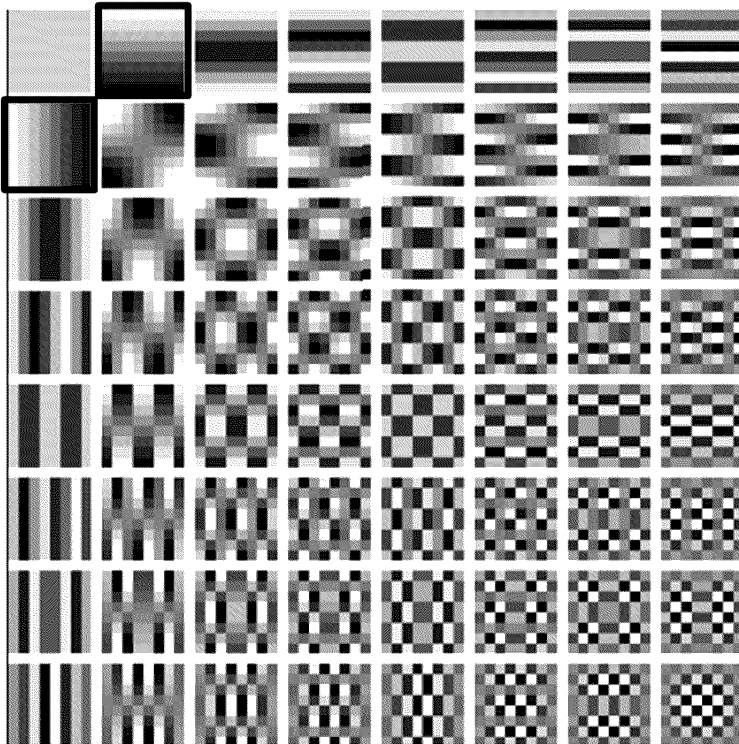
(Synchronous DRAM, SDRAM), 2배속 SDRAM(Double Data Rate SDRAM, DDR SDRAM), 증강형 SDRAM(Enhanced SDRAM, ESDRAM), 동기화 DRAM(SyncLink DRAM, SLDRAM) 및 직접 램버스 램(Direct Rambus RAM, DRRAM) 과 같은 다양한 RAM을 의미한다.

[0104]

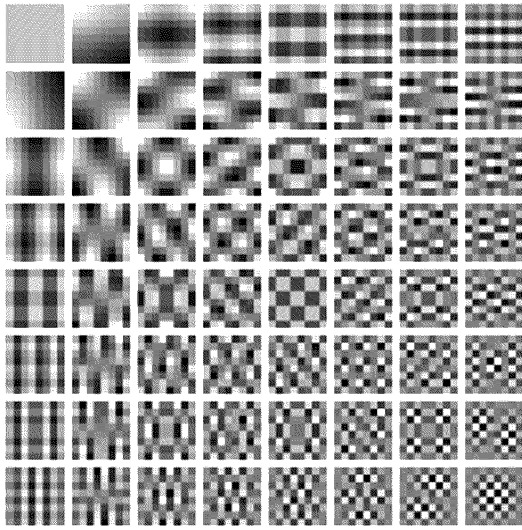
본 실시례 및 본 명세서에 첨부된 도면은 전술한 기술에 포함되는 기술적 사상의 일부를 명확하게 나타내고 있는 것에 불과하며, 전술한 기술의 명세서 및 도면에 포함된 기술적 사상의 범위 내에서 당업자가 용이하게 유추할 수 있는 변형 예와 구체적인 실시례는 모두 전술한 기술의 권리범위에 포함되는 것이 자명하다고 할 것이다.

도면

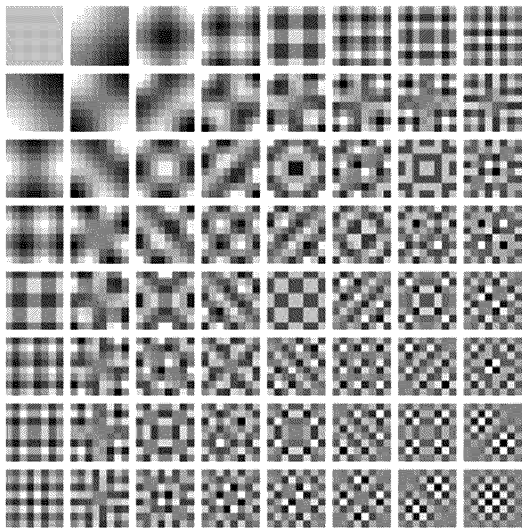
도면1



도면2

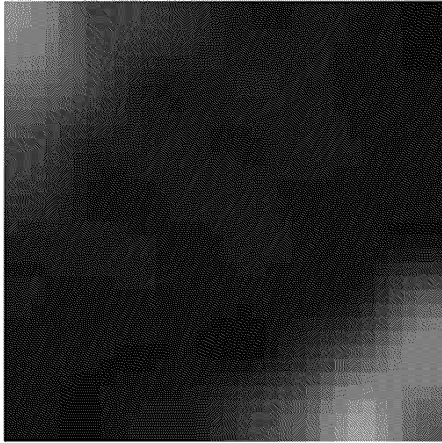


(A)

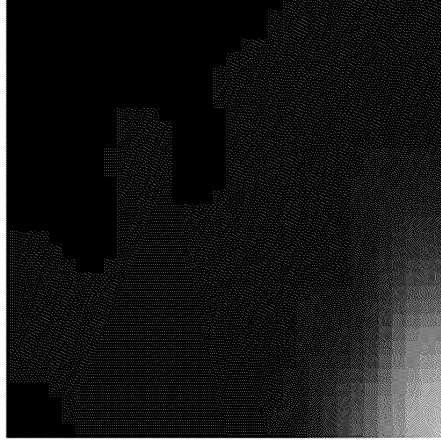


(B)

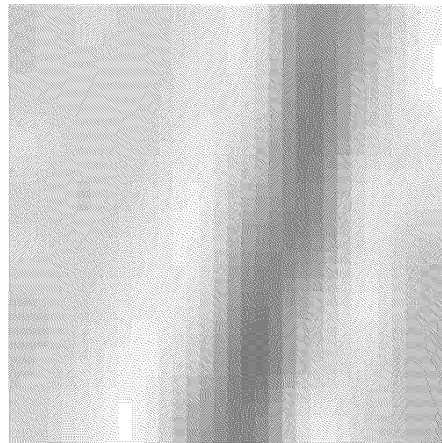
도면3



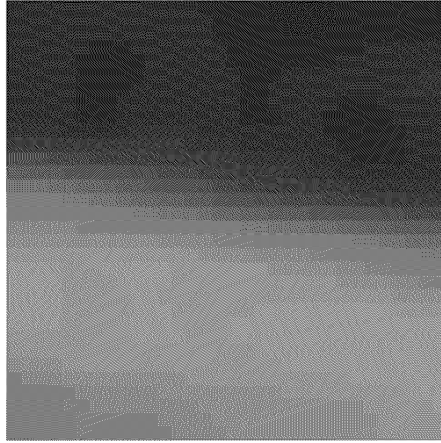
(A)



(B)

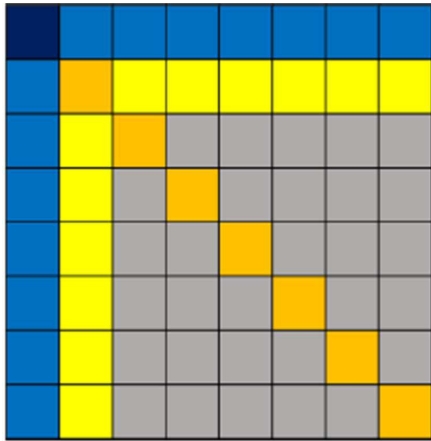


(C)

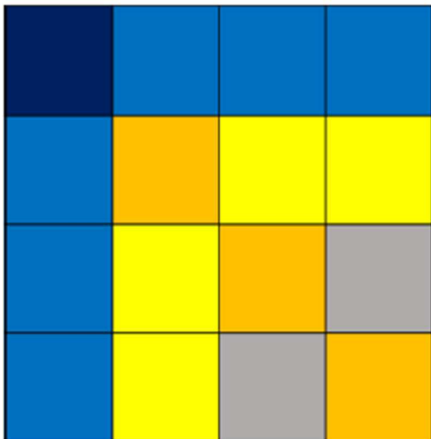


(D)

도면4

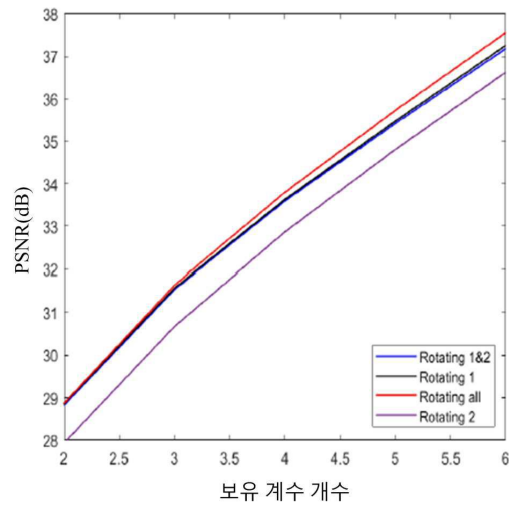


(A)

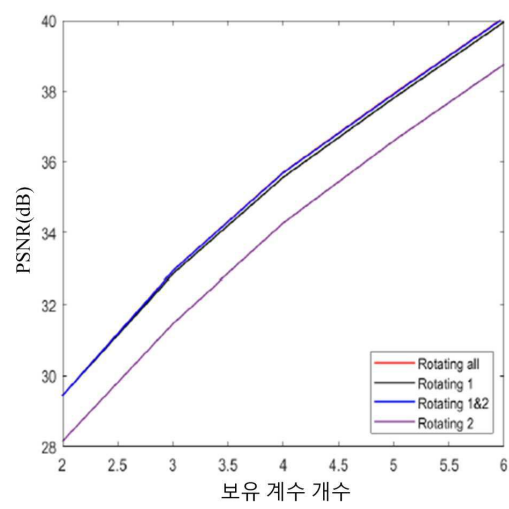


(B)

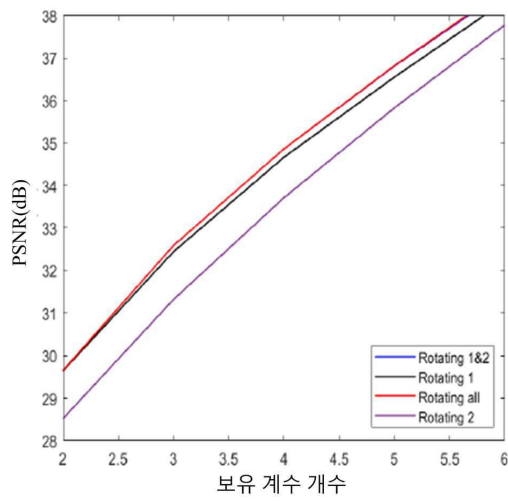
도면5



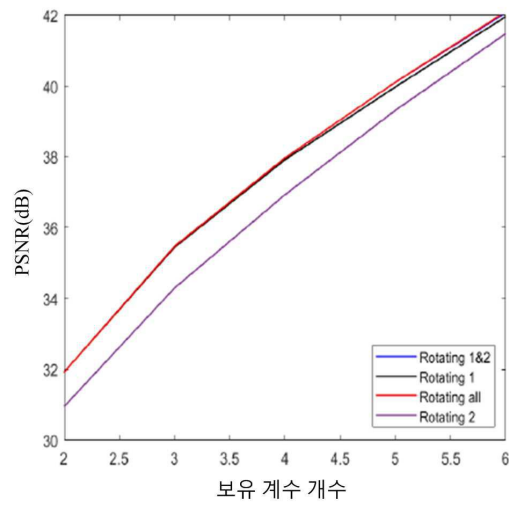
(A)



(B)

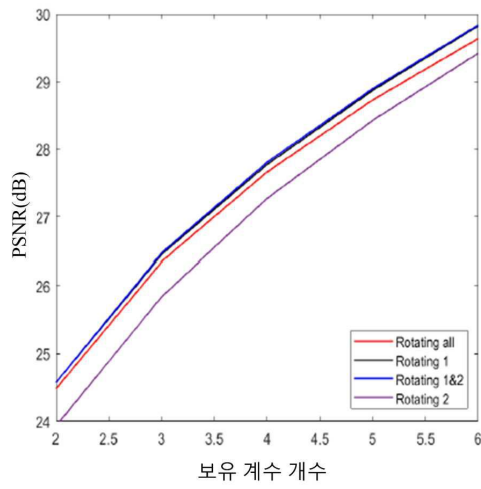


(C)

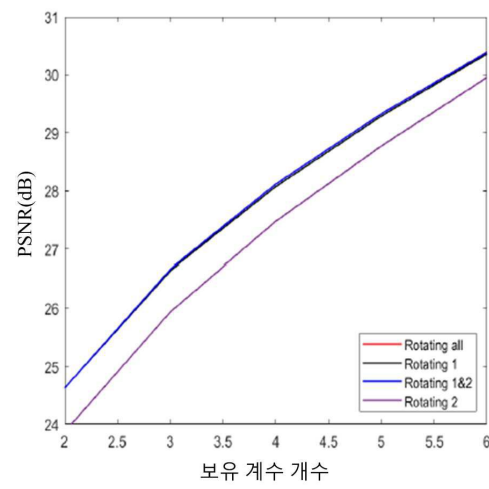


(D)

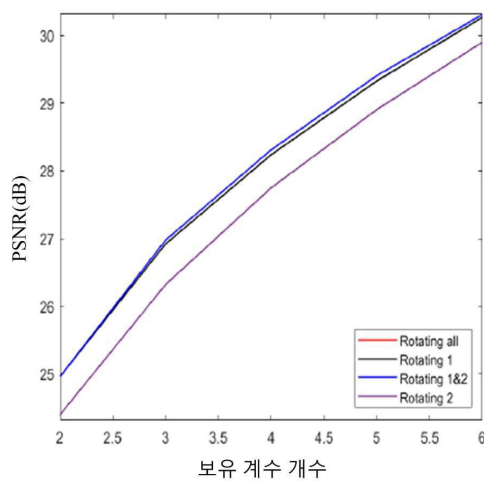
도면6



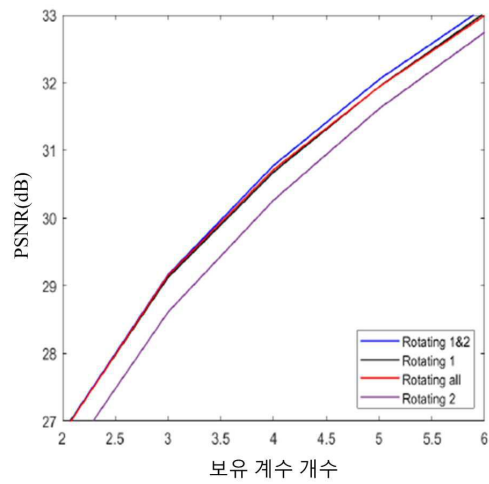
(A)



(B)



(C)



(D)

도면7

$$R(\theta) = \begin{bmatrix} 1 & & & & & & & & & \\ & \cos \theta & & \sin \theta & & & & & & \\ & & \cos \theta & & & & & & & \\ & & & \sin \theta & & & & & & \\ & & & & 1 & & & & & \\ & & & & & \cos \theta & & \sin \theta & & \\ & & & & & & \cos \theta & & \sin \theta & \\ & & & & & & & \sin \theta & & \\ & & & & & & & & 1 & \\ & & & & & & & & & \cos \theta & \\ & & & & & & & & & & \sin \theta & \\ & & & & & & & & & & & \cos \theta & \\ & & & & & & & & & & & & 1 \end{bmatrix}$$

도면8

