



(19) 대한민국특허청(KR)
(12) 등록특허공보(B1)

(45) 공고일자 2013년12월09일
(11) 등록번호 10-1340086
(24) 등록일자 2013년12월03일

(51) 국제특허분류(Int. Cl.)

H04N 13/00 (2006.01)

(21) 출원번호 10-2012-0123131

(22) 출원일자 2012년11월01일

심사청구일자 2012년11월01일

(56) 선행기술조사문헌

KR1020120017658 A

KR1020090090148 A

KR1020120042323 A

(73) 특허권자

서강대학교산학협력단

서울특별시 마포구 백범로 35 (신수동, 서강대학교)

엠텍비전 주식회사

경기도 성남시 분당구 판교로 344, 10층(삼평동, 엠텍아이티타워)

(72) 발명자

김영곤

서울시 성동구 송정동 96-23호

박래홍

서울시 영등포구 문래동 3가 54번지 문래자이 아파트 112-1603호

(74) 대리인

한상천

전체 청구항 수 : 총 17 항

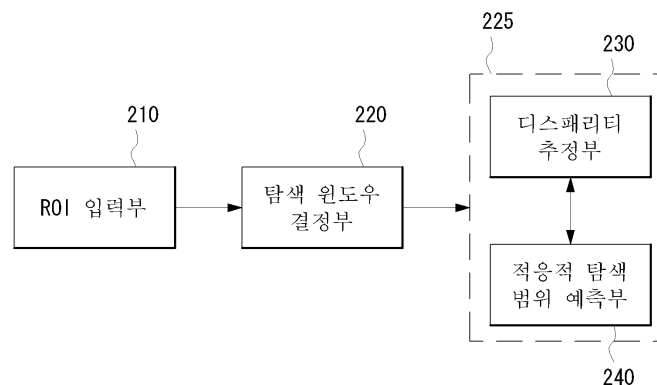
심사관 : 김기호

(54) 발명의 명칭 관심 영역의 자동 초점을 위해 동영상에서의 적응적 탐색 범위 예측을 이용한 디스패리티 추정 방법 및 장치

(57) 요약

관심 영역의 자동 초점을 위한 동영상에서의 적응적 탐색 범위 예측을 이용한 디스패리티 추정 방법 및 장치가 개시된다. 디스패리티 추정 장치는, 입력된 관심 영역(ROI)의 좌표 정보를 이용하여 디스패리티(disparity) 추정을 위한 탐색 윈도우를 결정하는 탐색 윈도우 결정부; 및 상기 결정된 탐색 윈도우 및 동영상의 템포럴(temporal) 정보를 이용하여 각 프레임별로 탐색 범위 결정 및 디스패리티 값 추정을 수행하는 디스패리티 근사화 유닛을 포함한다.

대표도 - 도2



특허청구의 범위

청구항 1

자동 초점(Auto convergence)을 위한 적응적 탐색 범위 예측을 이용한 디스패리티 추정 장치에 있어서,

입력된 관심 영역(ROI)의 좌표 정보를 이용하여 디스패리티(disparity) 추정을 위한 탐색 윈도우를 결정하는 탐색 윈도우 결정부; 및

상기 결정된 탐색 윈도우 및 동영상의 템포럴(temporal) 정보를 이용하여 각 프레임별로 탐색 범위 결정 및 디스패리티 값 추정을 수행하는 디스패리티 근사화 유닛을 포함하는 디스패리티 추정 장치.

청구항 2

제1항에 있어서,

상기 템포럴 정보는 시간적으로 앞선 프레임에 대해 추정된 디스패리티 값인 것을 특징으로 하는 디스패리티 추정 장치.

청구항 3

제1항에 있어서,

상기 탐색 윈도우는 하기 수학적식

$$\Omega_n = \left\{ (i, j) \mid x_n^1 \leq i \leq \left\lfloor \frac{x_n^1 + x_n^2}{2} \right\rfloor, \left\lfloor \frac{y_n^1 + y_n^2}{2} \right\rfloor - \frac{H}{2} \leq j \leq \left\lfloor \frac{y_n^1 + y_n^2}{2} \right\rfloor + \frac{H}{2} \right\}$$

에 의해 결정되되,

(x_n^1, y_n^1) 및 (x_n^2, y_n^2) 는 각각 n번째 프레임의 관심 영역 내에 위치한 좌측 상단 모서리의 좌표와 우측 하단 모서리의 좌표이고, i와 j는 각각 x좌표와 y좌표를 나타내며, H는 짝수로 정의된 윈도우의 높이이고, $\lfloor x \rfloor$ 는 x를 넘지 않는 최소 정수를 나타내며, A|B는 B 조건을 만족하는 A를 나타내는 것을 특징으로 하는 디스패리티 추정 장치.

청구항 4

제1항에 있어서,

각 프레임별로 개별 결정되는 상기 디스패리티 값 중 n번째 프레임의 디스패리티 값($\hat{d}_n$)은 수학적식 $\hat{d}_n = \tilde{d}_n + \Delta \hat{d}_n$ 에 의해 추정되되,

상기 $\tilde{d}_n$ 은 n번째 프레임의 디스패리티 근사값이고, 상기 $\Delta \hat{d}_n$ 은 n번째 프레임의 디스패리티 보상값인 것을 특징으로 하는 디스패리티 추정 장치.

청구항 5

제4항에 있어서,

n번째 프레임의 상기 디스패리티 근사값($\tilde{d}_n$)은 수학식 $\tilde{d}_n = \hat{d}_{n-1} + d'_{n-1}$ 에 의해 산출되되,

상기 d'_{n-1} 은 n-1 번째 프레임에서 추정된 디스패리티 값과 n-2번째 프레임에서 추정된 디스패리티 값 간의 차이 값인 것을 특징으로 하는 디스패리티 추정 장치.

청구항 6

제4항에 있어서,

n번째 프레임의 상기 디스패리티 보상값($\Delta\hat{d}_n$)은 하기 수학식

$$\Delta\hat{d}_n = \underset{\Delta d_n}{\operatorname{argmin}} \sum_{(p,q) \in \Omega_n} |I_L(p + \tilde{d}_n, q) - I_R(p + \tilde{d}_n + \Delta d_n, q)|, \quad -D_n \leq \Delta\hat{d}_n \leq D_n, 1 \leq n < N$$

에 의해 산출되되,

I_L 과 I_R 은 2차원 평면으로 투사된 좌측 및 우측 2차원 영상이고, (p, q)는 탐색 윈도우 내의 픽셀의 좌표이며, N은 동영상의 총 프레임 수이고, D_n 은 산출된 탐색 범위의 값인 것을 특징으로 하는 디스패리티 추정 장치.

청구항 7

제6항에 있어서,

각 프레임별로 개별 결정되는 상기 탐색 범위의 값(D_n)은 하기 수학식

$$D_n = D^0 + \left\lfloor \left| \tilde{D}_n \right| \right\rfloor$$

에 의해 결정되되,

상기 D^0 는 미리 지정된 오프셋(offset)이고, 상기 $\tilde{D}_n$ 은 k(임의의 자연수)개의 디스패리티 값들의 각 차이들에 대한 평균값이고, $\left\lfloor \left| \tilde{D}_n \right| \right\rfloor$ 은 $\tilde{D}_n$ 의 절대값을 넘지 않는 최소 정수를 나타내는 것을 특징으로 하는 디스패리티 추정 장치.

청구항 8

제7항에 있어서,

상기 k는 3인 것을 특징으로 하는 디스패리티 추정 장치.

청구항 9

디스패리티 추정 장치에서 수행되는, 자동 초점(Auto convergence)을 위한 적응적 탐색 범위 예측을 이용한 디스패리티 추정 방법에 있어서,

입력된 관심 영역(ROI)의 좌표 정보를 이용하여 디스패리티(disparity) 추정을 위한 탐색 윈도우를 결정하는 단

계; 및

상기 결정된 탐색 윈도우 및 동영상의 템포럴(temporal) 정보를 이용하여 각 프레임별로 탐색 범위 결정 및 디스패리티 값 추정을 수행하는 단계를 포함하는 디스패리티 추정 방법.

청구항 10

제9항에 있어서,

상기 템포럴 정보는 시간적으로 앞선 프레임에 대해 추정된 디스패리티 값인 것을 특징으로 하는 디스패리티 추정 방법.

청구항 11

제9항에 있어서,

상기 탐색 윈도우는 하기 수학적식

$$\Omega_n = \left\{ (i, j) \mid x_n^1 \leq i \leq \left\lfloor \frac{x_n^1 + x_n^2}{2} \right\rfloor, \left\lfloor \frac{y_n^1 + y_n^2}{2} \right\rfloor - \frac{H}{2} \leq j \leq \left\lfloor \frac{y_n^1 + y_n^2}{2} \right\rfloor + \frac{H}{2} \right\}$$

에 의해 결정되되,

(x_n^1, y_n^1) 및 (x_n^2, y_n^2) 는 각각 n번째 프레임의 관심 영역 내에 위치한 좌측 상단 모서리의 좌표와 우측 하단 모서리의 좌표이고, i와 j는 각각 x좌표와 y좌표를 나타내며, H는 짝수로 정의된 윈도우의 높이이고, $\lfloor x \rfloor$ 는 x를 넘지 않는 최소 정수를 나타내며, A|B는 B 조건을 만족하는 A를 나타내는 것을 특징으로 하는 디스패리티 추정 방법.

청구항 12

제9항에 있어서,

각 프레임별로 개별 결정되는 상기 디스패리티 값 중 n번째 프레임의 디스패리티 값($\hat{d}_n$)은 수학적식 $\hat{d}_n = \tilde{d}_n + \Delta \hat{d}_n$ 에 의해 추정되되,

상기 $\tilde{d}_n$ 은 n번째 프레임의 디스패리티 근사값이고, 상기 $\Delta \hat{d}_n$ 은 n번째 프레임의 디스패리티 보상값인 것을 특징으로 하는 디스패리티 추정 방법.

청구항 13

제12항에 있어서,

n번째 프레임의 상기 디스패리티 근사값($\tilde{d}_n$)은 수학적식 $\tilde{d}_n = \hat{d}_{n-1} + d'_{n-1}$ 에 의해 산출되되,

상기 d'_{n-1} 은 n-1 번째 프레임에서 추정된 디스패리티 값과 n-2번째 프레임에서 추정된 디스패리티 값 간의 차이 값인 것을 특징으로 하는 디스패리티 추정 방법.

청구항 14

제12항에 있어서,

n번째 프레임의 상기 디스패리티 보상값($\Delta \hat{d}_n$)은 하기 수학적식

$$\Delta \hat{d}_n = \underset{\Delta d_n}{\operatorname{argmin}} \sum_{(p,q) \in \Omega_n} \left| I_L(p + \tilde{d}_n, q) - I_R(p + \tilde{d}_n + \Delta d_n, q) \right|, \quad -D_n \leq \Delta \hat{d}_n \leq D_n, 1 \leq n < N$$

에 의해 산출되되,

I_L 과 I_R 은 2차원 평면으로 투사된 좌측 및 우측 2차원 영상이고, (p, q)는 탐색 윈도우 내의 픽셀의 좌표이며, N은 동영상의 총 프레임 수이고, D_n 은 산출된 탐색 범위의 값인 것을 특징으로 하는 디스패리티 추정 방법.

청구항 15

제14항에 있어서,

각 프레임별로 개별 결정되는 상기 탐색 범위의 값(D_n)은 하기 수학적식

$$D_n = D^0 + \left\lfloor \left| \tilde{D}_n \right| \right\rfloor$$

에 의해 결정되되,

상기 D^0 는 미리 지정된 오프셋(offset)이고, 상기 $\tilde{D}_n$ 은 k(임의의 자연수)개의 디스패리티 값들의 각 차이들에 대한 평균값이고, $\left\lfloor \left| \tilde{D}_n \right| \right\rfloor$ 은 $\tilde{D}_n$ 의 절대값을 넘지 않는 최소 정수를 나타내는 것을 특징으로 하는 디스패리티 추정 방법.

청구항 16

제15항에 있어서,

상기 k는 3인 것을 특징으로 하는 디스패리티 추정 방법.

청구항 17

제9항 내지 제16항 중 어느 한 항에 기재된 디스패리티 추정 방법을 수행하기 위하여 디지털 처리 장치에 의해 실행될 수 있는 명령어들의 프로그램이 기록된 기록매체.

명세서

기술분야

본 발명은 관심 영역의 자동 초점을 위해 동영상에서의 적응적 탐색 범위 예측을 이용한 디스패리티 추정 방법 및 장치에 관한 것이다.

배경기술

[0001]

- [0002] 최근 3차원 영상 처리에 대한 연구가 활발하게 진행되고 있으며, 그 결과 3차원 효과를 구현할 수 있는 다양한 장치들이 개발되고 있다.
- [0003] 3차원 디스플레이 장치를 이용하여 3차원 이미지 효과를 느끼도록 하는 방식들은 안경 방식과 무안경 방식으로 구분될 수 있다.
- [0004] 안경 방식으로는 셔터 안경 방식과 편광 안경 방식이 있고, 셔터 안경 방식은 디스플레이 장치에서 좌측 뷰(view)와 우측 뷰를 차례대로 좌안(왼쪽 눈)과 우안(오른쪽 눈)에서 동기화를 맞추어 볼 수 있도록 뷰를 번갈아 차단하는 방식이고, 편광 안경 방식은 편광 필터를 이용하여 좌측 뷰와 우측 뷰에서 원하는 색만이 투과되도록 하여 3차원 효과를 내도록 하는 방식이다.
- [0005] 또한, 안경 없이 디스플레이 장치만으로 3차원 효과를 느낄 수 있도록 하는 무안경 방식으로 패럴랙스 배리어(parallax barrier) 방식과 렌티큘러(lenticular) 방식이 있다. 패럴랙스 배리어 방식은 좌안과 우안의 영상을 각각의 눈에 해당하는 영상만 보여주되 해당하지 않는 영상은 차단하는 방식으로, 좌측 뷰는 좌안에서만 우측 뷰는 우안에서만 볼 수 있도록 차단막을 적절히 사용하여 시차를 분리하는 방식이다.
- [0006] 일반적으로 3차원 상에 존재하는 물체의 한 지점에 대해 좌/우 영상으로 투사(projection)된 좌표 차이를 디스패리티(disparity) 또는 패럴랙스(parallax)라고 하는데, 두 카메라의 광축이 한 점에서 만날 때 그 점에 위치한 물체의 디스패리티는 0(zero)이며, 그 지점을 수렴점(convergence point)이라고 한다.
- [0007] 관심 영역(ROI)의 디스패리티가 음의 값을 가지면 음시차(negative parallax)라고 하고, 양의 값을 가지면 양시차(positive parallax)라고 정의된다. 디스플레이 장치를 기준할 때, 수렴점에 위치한 물체(즉 디스패리티가 0인 지점의 물체)는 3차원 공간 상에서 디스플레이 상에 위치하지만, 음시차를 가지는 물체는 3차원 공간상에서 디스플레이 앞에 위치하고, 양시차를 가지는 물체는 디스플레이 뒤에 위치한다.
- [0008] 양시차는 음시차든 그 크기가 큰 값을 가지는 물체는 사용자가 3차원 영상으로 시청할 때 눈에 쉽게 피로감을 느끼게 한다. 이는, 큰 값의 양시차 또는 음시차를 가지는 물체에 초점을 맞추기 위해서는 두 눈의 초점을 디스플레이보다 앞 또는 뒤에 맞추어야 하기 때문이다.
- [0009] 따라서, 관심 영역에 수렴점을 맞추는 것은 두 눈의 피로를 줄이고 편안함을 주면서 동시에 수렴점 근처 물체들의 상대적인 디스패리티 변화로 인해 3차원 효과를 느끼게 해주기 때문에 반드시 필요하며, 관심 영역에 수렴점을 맞추기 위한 다양한 방법들이 소개되고 있다.
- [0010] 관심 영역에 수렴점을 맞추기 위한 하나의 방법은 카메라의 렌즈나 CCD(Charge Coupled Device) 센서를 하드웨어적으로 수평 방향 이동시켜 관심 영역에 초점이 맞도록 좌측 및 우측 영상을 취득하거나, 두 카메라의 광축이 관심 영역에 교차되도록 두 카메라의 각도를 변화시키는 방법이 그것이다.
- [0011] 그리고 다른 방법은, 관심 영역에 존재하는 물체의 추정된 디스패리티 크기만큼 두 영상을 수평 방향으로 이동시키는 방법이다. 이 방법을 적용하기 위해서는 전처리 과정으로 두 영상간에 정렬(rectification)이 이루어져야 하며, 만약 정렬이 이루어지지 않은 두 영상을 사용하는 경우에는 결과 영상에서 경계 부분에 키스톤 왜곡(keystone distortion)이 발생하여 두 눈의 피로를 가중시키는 원인이 된다.
- [0012] 그러나, 전술한 방법들은 사용자가 관심 있게 보는 물체와 무관하게 모든 프레임에 대해서 고정된 디스패리티 값 만큼만 적용되어 관심 영역에 수렴점이 맞지 않고, 큰 양시차 또는 음시차를 가질 수 밖에 없는 한계가 있기 때문에 두 눈이 쉽게 피로를 느끼게 되는 문제점이 있다.

발명의 내용

해결하려는 과제

- [0013] 본 발명은 이전에 추정된 디스패리티(disparity)로부터 1차 테일러 급수(first-order Taylor series)와 적응적인 탐색범위 예측을 사용하여 현재 프레임의 디스패리티를 효율적이고 정확하게 추정할 수 있어 동영상에서 모든 프레임의 관심 영역을 수렴점(convergence point)에 맞출 수 있는 관심 영역의 자동 초점을 위한 동영상에서의 적응적 탐색 범위 예측을 이용한 디스패리티 추정 방법 및 장치를 제공하기 위한 것이다.

[0014] 본 발명의 이외의 목적들은 하기의 설명을 통해 쉽게 이해될 수 있을 것이다.

과제의 해결 수단

[0015] 본 발명의 일 측면에 따르면, 자동 초점(Auto convergence)을 위한 적응적 탐색 범위 예측을 이용한 디스패리티 추정 장치에 있어서, 입력된 관심 영역(ROI)의 좌표 정보를 이용하여 디스패리티(disparity) 추정을 위한 탐색 윈도우를 결정하는 탐색 윈도우 결정부; 및 상기 결정된 탐색 윈도우 및 동영상의 템포럴(temporal) 정보를 이용하여 각 프레임별로 탐색 범위 결정 및 디스패리티 값 추정을 수행하는 디스패리티 근사화 유닛을 포함하는 디스패리티 추정 장치가 제공된다.

[0016] 상기 템포럴 정보는 시간적으로 앞선 프레임에 대해 추정된 디스패리티 값일 수 있다.

[0017] 상기 탐색 윈도우는 하기 수학적식

$$\Omega_n = \left\{ (i, j) \mid x_n^1 \leq i \leq \left\lfloor \frac{x_n^1 + x_n^2}{2} \right\rfloor, \left\lfloor \frac{y_n^1 + y_n^2}{2} \right\rfloor - \frac{H}{2} \leq j \leq \left\lfloor \frac{y_n^1 + y_n^2}{2} \right\rfloor + \frac{H}{2} \right\}$$

[0019] 에 의해 결정되되, (x_n^1, y_n^1) 및 (x_n^2, y_n^2) 는 각각 n번째 프레임의 관심 영역 내에 위치한 좌측 상단 모서리의 좌표와 우측 하단 모서리의 좌표이고, H는 짝수로 정의된 윈도우의 높이이며, $\lfloor x \rfloor$ 는 x를 넘지 않는 최소 정수를 의미할 수 있다.

[0020] 각 프레임별로 개별 결정되는 상기 디스패리티 값 중 n번째 프레임의 디스패리티 값($\hat{d}_n$)은 수학적식 $\hat{d}_n = \tilde{d}_n + \Delta \hat{d}_n$ 에 의해 추정되되, 상기 $\tilde{d}_n$ 은 n번째 프레임의 디스패리티 근사값이고, 상기 $\Delta \hat{d}_n$ 은 n번째 프레임의 디스패리티 보상값일 수 있다.

[0021] n번째 프레임의 상기 디스패리티 근사값($\tilde{d}_n$)은 수학적식 $\tilde{d}_n = \hat{d}_{n-1} + d'_{n-1}$ 에 의해 산출되되, 상기 d'_{n-1} 은 n-1번째 프레임에서 추정된 디스패리티 값과 n-2번째 프레임에서 추정된 디스패리티 값 간의 차이 값일 수 있다.

[0022] n번째 프레임의 상기 디스패리티 보상값($\Delta \hat{d}_n$)은 하기 수학적식

$$\Delta \hat{d}_n = \underset{\Delta \hat{d}_n}{\operatorname{argmin}} \sum_{(p, q) \in \Omega_n} \left| I_L(p + \tilde{d}_n, q) - I_R(p + \tilde{d}_n + \Delta \hat{d}_n, q) \right|, \quad -D_n \leq \Delta \hat{d}_n \leq D_n, 1 \leq n < N$$

[0024] 에 의해 산출되되, I_L 과 I_R 은 2차원 평면으로 투사된 좌측 및 우측 2차원 영상이고, (p, q)는 탐색 윈도우 내의 픽셀의 좌표이며, N은 동영상의 총 프레임 수이고, D_n 은 산출된 탐색 범위의 값일 수 있다.

[0025] 각 프레임별로 개별 결정되는 상기 탐색 범위의 값(D_n)은 하기 수학적식

$$D_n = D^0 + \left\lfloor \left| \tilde{D}_n \right| \right\rfloor$$

[0026] 에 의해 결정되되, 상기 D^0 는 미리 지정된 오프셋(offset)이고, 상기 $\tilde{D}_n$ 은 k(임의의 자연수)개의 디스패리티 값들의 각 차이들에 대한 평균값이고, $\left\lfloor \left| \tilde{D}_n \right| \right\rfloor$ 은 $\tilde{D}_n$ 의 절대값을 넘지 않는 최소 정수를 나타내는 것일 수 있다. 상기 k는 3일 수 있다.

[0027] 본 발명의 다른 측면에 따르면, 디스패리티 추정 장치에서 수행되는, 자동 초점(Auto convergence)을 위한 적응적 탐색 범위 예측을 이용한 디스패리티 추정 방법에 있어서, 입력된 관심 영역(ROI)의 좌표 정보를 이용하여 디스패리티(disparity) 추정을 위한 탐색 윈도우를 결정하는 단계; 및 상기 결정된 탐색 윈도우 및 동영상의 템

포털(temporal) 정보를 이용하여 각 프레임별로 탐색 범위 결정 및 디스패리티 값 추정을 수행하는 단계를 포함하는 디스패리티 추정 방법이 제공된다.

[0028] 상기 템포럴 정보는 시간적으로 앞선 프레임에 대해 추정된 디스패리티 값일 수 있다.

[0029] 상기 탐색 윈도우는 하기 수학적

$$\Omega_n = \left\{ (i, j) \mid x_n^1 \leq i \leq \left\lfloor \frac{x_n^1 + x_n^2}{2} \right\rfloor, \left\lfloor \frac{y_n^1 + y_n^2}{2} \right\rfloor - \frac{H}{2} \leq j \leq \left\lfloor \frac{y_n^1 + y_n^2}{2} \right\rfloor + \frac{H}{2} \right\}$$

[0031] 에 의해 결정되되, (x_n^1, y_n^1) 및 (x_n^2, y_n^2) 는 각각 n번째 프레임의 관심 영역 내에 위치한 좌측 상단 모서리의 좌표와 우측 하단 모서리의 좌표이고, H는 짝수로 정의된 윈도우의 높이이며, $\lfloor x \rfloor$ 는 x를 넘지 않는 최소 정수를 나타내는 것일 수 있다.

[0032] 각 프레임별로 개별 결정되는 상기 디스패리티 값 중 n번째 프레임의 디스패리티 값($\hat{d}_n$)은 수학적 $\hat{d}_n = \tilde{d}_n + \Delta \hat{d}_n$ 에 의해 추정되되, 상기 $\tilde{d}_n$ 은 n번째 프레임의 디스패리티 근사값이고, 상기 $\Delta \hat{d}_n$ 은 n번째 프레임의 디스패리티 보상값일 수 있다.

[0033] n번째 프레임의 상기 디스패리티 근사값($\tilde{d}_n$)은 수학적 $\tilde{d}_n = \hat{d}_{n-1} + d'_{n-1}$ 에 의해 산출되되, 상기 d'_{n-1} 은 n-1번째 프레임에서 추정된 디스패리티 값과 n-2번째 프레임에서 추정된 디스패리티 값 간의 차이 값일 수 있다.

[0034] n번째 프레임의 상기 디스패리티 보상값($\Delta \hat{d}_n$)은 하기 수학적

$$\Delta \hat{d}_n = \underset{\Delta d_n}{\operatorname{argmin}} \sum_{(p, q) \in \Omega_n} |I_L(p + \tilde{d}_n, q) - I_R(p + \tilde{d}_n + \Delta d_n, q)|, \quad -D_n \leq \Delta \hat{d}_n \leq D_n, 1 \leq n < N$$

[0036] 에 의해 산출되되, I_L 과 I_R 은 2차원 평면으로 투사된 좌측 및 우측 2차원 영상이고, (p, q)는 탐색 윈도우 내의 픽셀의 좌표이며, N은 동영상의 총 프레임 수이고, D_n 은 산출된 탐색 범위의 값일 수 있다.

[0037] 각 프레임별로 개별 결정되는 상기 탐색 범위의 값(D_n)은 하기 수학적

$$D_n = D^0 + \left\lfloor \left| \tilde{D}_n \right| \right\rfloor$$

[0038] 에 의해 결정되되, 상기 D^0 는 미리 지정된 오프셋(offset)이고, 상기 $\tilde{D}_n$ 은 k(임의의 자연수)개의 디스패리티 값들의 각 차이들에 대한 평균값이고, $\left\lfloor \left| \tilde{D}_n \right| \right\rfloor$ 은 $\tilde{D}_n$ 의 절대값을 넘지 않는 최소 정수를 나타내는 것일 수 있다. 상기 k는 3일 수 있다.

[0039] 전술한 것 외의 다른 측면, 특징, 이점이 이하의 도면, 특허청구범위 및 발명의 상세한 설명으로부터 명확해질 것이다.

발명의 효과

[0040] 본 발명의 실시예에 따르면, 이전에 추정된 디스패리티(disparity)로부터 1차 테일러 급수(first-order Taylor series)와 적응적인 탐색범위 예측을 사용하여 현재 프레임의 디스패리티를 효율적이고 정확하게 추정할 수 있어 동영상에서 모든 프레임의 관심 영역을 수렴점(convergence point)에 맞추도록 하는 효과가 있다.

도면의 간단한 설명

- [0041] 도 1은 수렴점(convergence point)를 맞추기 위한 원근 기하(perspective geometry)를 나타낸 도면.
 도 2는 본 발명의 일 실시예에 따른 디스패리티(disparity) 추정 장치의 구성을 개략적으로 나타낸 도면.
 도 3은 본 발명의 일 실시예에 따른 n번째 프레임에서의 디스패리티 추정 방법을 나타낸 그래프.
 도 4는 본 발명의 일 실시예에 따른 디스패리티 추정 방법을 나타낸 순서도.
 도 5a 내지 도 5f는 본 발명의 일 실시예에 따른 디스패리티 추정 장치의 효과를 설명하기 위한 도면.

발명을 실시하기 위한 구체적인 내용

- [0042] 본 발명은 다양한 변경을 가할 수 있고 여러 가지 실시예를 가질 수 있는 바, 특정 실시예들을 도면에 예시하고 상세한 설명에 상세하게 설명하고자 한다. 그러나 이는 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다.
- [0043] 어떤 구성요소가 다른 구성요소에 "연결되어" 있다거나 "접속되어" 있다고 언급된 때에는, 그 다른 구성요소에 직접적으로 연결되어 있거나 또는 접속되어 있을 수도 있지만, 중간에 다른 구성요소가 존재할 수도 있다고 이해되어야 할 것이다. 반면에, 어떤 구성요소가 다른 구성요소에 "직접 연결되어" 있다거나 "직접 접속되어" 있다고 언급된 때에는, 중간에 다른 구성요소가 존재하지 않는 것으로 이해되어야 할 것이다.
- [0044] 본 명세서에서 사용한 용어는 단지 특정한 실시예를 설명하기 위해 사용된 것으로, 본 발명을 한정하려는 의도가 아니다. 단수의 표현은 문맥상 명백하게 다르게 뜻하지 않는 한, 복수의 표현을 포함한다. 본 명세서에서, "포함하다" 또는 "가지다" 등의 용어는 명세서상에 기재된 특징, 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것이 존재함을 지정하려는 것이지, 하나 또는 그 이상의 다른 특징들이나 숫자, 단계, 동작, 구성요소, 부품 또는 이들을 조합한 것들의 존재 또는 부가 가능성을 미리 배제하지 않는 것으로 이해되어야 한다.
- [0045] 제1, 제2 등의 용어는 다양한 구성요소들을 설명하는데 사용될 수 있지만, 상기 구성요소들은 상기 용어들에 의해 한정되어서는 안 된다. 상기 용어들은 하나의 구성요소를 다른 구성요소로부터 구별하는 목적으로만 사용된다.
- [0046] 또한, 명세서에 기재된 "...부", "...유닛", "...모듈" 등의 용어는 적어도 하나의 기능이나 동작을 처리하는 단위를 의미하며, 이는 하드웨어나 소프트웨어 또는 하드웨어 및 소프트웨어의 결합으로 구현될 수 있다.
- [0047] 또한, 첨부 도면을 참조하여 설명함에 있어, 도면 부호에 관계없이 동일한 구성 요소는 동일한 참조부호를 부여하고 이에 대한 중복되는 설명은 생략하기로 한다. 본 발명을 설명함에 있어서 관련된 공지 기술에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.
- [0048] 도 1은 수렴점(convergence point)를 맞추기 위한 원근 기하(perspective geometry)를 나타낸 도면이다.
- [0049] 앞서 설명한 수렴점을 맞추기 위한 각 방법들은 모두 원근 기하라는 이론적 배경을 바탕으로 한다. 원근 기하의 개념을 이용하면, 2차원(two-dimensional) 영상 평면상의 한 점이 3차원 월드(world) 좌표계로 역투사(back-projection)된다. 스테레오(Stereo) 영상들을 정렬(rectification)하거나 수렴점을 관심 영역(ROI)에 맞추기 위해 역투사된 3차원 월드 좌표계 상의 점들로부터 가상 뷰(view)를 만들어 낸다.
- [0050] 도 1은 Y_w 축을 기준으로 X_w 축과 Z_w 축에 대한 월드 좌표계를 나타내며, 정렬된 두 영상 및 동차 좌표(homogeneous coordinate)를 사용하여 한 점 X_w 에서 수렴점을 맞추는 원근 기하를 나타낸다. 월드 좌표계 내의 한 점 X_w 는 동차 좌표(homogeneous coordinate)를 사용하여 $X_w = [X_w, Y_w, Z_w, 1]^T$ 로 정의되고, 동차 좌표계를 사용하여 월드 좌표계에서의 한 점 X_w 가 2차원 평면으로 투사되는 x 는 하기 수학식 1로 정의할 수 있다.

수학식 1

$$\mathbf{x} = K[I | \mathbf{0}] \begin{bmatrix} R & -R\mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \mathbf{X}_w = K[I | \mathbf{0}] \mathbf{X}_c$$

[0051]

[0052]

여기서, K는 3x3 카메라 내부 파라미터로서, 카메라 좌표계에서의 한 점을 2차원 평면으로 투사할 때 사용된다. R과 t는 카메라 외부 파라미터로서, 월드 좌표계에서 한 점을 카메라 좌표계로 투사할 때 사용된다. R은 평행하게 놓인 카메라의 광 축 및 카메라의 중심과 Xw를 잇는 선과의 각도만큼 회전되도록 하는 3x3행렬이고, t는 월드 좌표계의 중심과 카메라 좌표계의 중심 사이에 이동한 크기와 방향을 3x1벡터로 나타낸다.

[0053]

월드 좌표계에서 좌측 카메라 좌표계와 우측 카메라 좌표계로 투사된 좌표 $\mathbf{X}_c^l$ 와 $\mathbf{X}_c^r$ 는 각각 하기 수학식 2와 3으로 표현될 수 있다.

수학식 2

$$\mathbf{X}_c^l = \begin{bmatrix} X_c^l \\ Y_c^l \\ Z_c^l \\ 1 \end{bmatrix} = \begin{bmatrix} R^l & -R^l\mathbf{t}^l \\ 0 & 1 \end{bmatrix} \mathbf{X}_w = \begin{bmatrix} \cos \theta & 0 & -\sin \theta & \frac{1}{2}b \cos \theta \\ 0 & 1 & 0 & 0 \\ \sin \theta & 0 & \cos \theta & \frac{1}{2}b \sin \theta \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

[0054]

수학식 3

$$\mathbf{X}_c^r = \begin{bmatrix} X_c^r \\ Y_c^r \\ Z_c^r \\ 1 \end{bmatrix} = \begin{bmatrix} R^r & -R^r\mathbf{t}^r \\ 0 & 1 \end{bmatrix} \mathbf{X}_w = \begin{bmatrix} \cos \theta & 0 & \sin \theta & -\frac{1}{2}b \cos \theta \\ 0 & 1 & 0 & 0 \\ -\sin \theta & 0 & \cos \theta & -\frac{1}{2}b \sin \theta \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

[0055]

[0056]

수학식 2 및 3과 가정 K=I를 적용하여 좌측 카메라 좌표 $\mathbf{X}_c^l$ 와 우측 카메라 좌표 $\mathbf{X}_c^r$ 가 좌측 2차원 평면과 우측 2차원 평면에 투사된 좌표 x^l 과 x^r 은 하기 수학식 4 및 5로 각각 표현될 수 있다.

[0057]

수학식 4

$$x^l = \frac{\left(X_w + \frac{b}{2}\right) \cos \theta - Z_w \sin \theta}{Z_w \cos \theta + \left(X_w + \frac{b}{2}\right) \sin \theta}, y^l = \frac{Y_w}{Z_w \cos \theta + \left(X_w + \frac{b}{2}\right) \sin \theta}$$

[0058]

수학식 5

$$x' = \begin{bmatrix} \left(X_w - \frac{b}{2}\right) \cos \theta + Z_w \sin \theta \\ Z_w \cos \theta - \left(X - \frac{b}{2}\right) \sin \theta \\ 1 \end{bmatrix}, y' = \begin{bmatrix} Y \\ Z \cos \theta - \left(X - \frac{b}{2}\right) \sin \theta \end{bmatrix}$$

[0059]

[0060] 2차원 평면에서 관심 영역의 디스패리티(d)를 구하는 식은 카메라 외부 파라미터 t에서 좌측 카메라 좌표 $\mathbf{X}_c^l$ 와 우측 카메라 좌표 $\mathbf{X}_c^r$ 만을 고려하기 때문에 하기 수학식 6과 같이 디스패리티(d)는 차원 평면으로 투사된 x축 좌표의 차이에 의해 연산될 수 있다.

수학식 6

$$d = x_l - x_r = \frac{\left(X + \frac{b}{2}\right) \cos \theta - Z \sin \theta}{Z \cos \theta + \left(X + \frac{b}{2}\right) \sin \theta} - \frac{\left(X - \frac{b}{2}\right) \cos \theta + Z \sin \theta}{Z \cos \theta - \left(X - \frac{b}{2}\right) \sin \theta}$$

[0061]

[0062] 여기서, 평행하게 놓인 두 카메라로부터 취득된 영상인 경우라면 θ 가 0이므로, 수학식 6은 하기 수학식 7과 같이 단순화될 수 있다.

수학식 7

$$d = \frac{b}{Z_w}$$

[0063]

[0064] 수학식 7을 참조하면, 디스패리티는 두 카메라간의 거리(base line) b 및 관심영역과 카메라 간의 거리 Z_w 에만 영향을 받음을 알 수 있다. 전술한 바와 같이, 관심 영역에 수렴점을 맞추기 위해서는 관심 영역의 디스패리티가 0이 되도록 하기 위해 관심 영역의 디스패리티만큼 영상을 수평방향으로 이동시킨다.

[0065] 수학식 7에 나타난 바와 같이, 월드 좌표계 상에서 움직이는 관심 영역의 Z_w 와 디스패리티(d)가 반비례 관계이므로, 관심 영역이 카메라에 접근할수록(즉 Z_w 가 작아질수록) 디스패리티가 커지고 그 반대의 경우 디스패리티가 작아진다.

[0066] 결과적으로, 디스패리티는 월드 좌표계에서의 한 점이 좌측과 우측 2차원 평면으로 투사된 두 x축 좌표의 차이를 나타내는데, 기준이 되는 좌측 2차원 영상에 대해 검출 알고리즘을 이용하여 x축 좌표를 알아내고, 이 좌표와 일치하는 우측 2차원 평면으로 투사된 또 다른 x축 좌표를 알아냄으로써 탐색범위 내에 존재하는 디스패리티를 알 수 있다.

[0067] 도 2는 본 발명의 일 실시예에 따른 디스패리티(disparity) 추정 장치의 구성을 개략적으로 나타낸 도면이고, 도 3은 본 발명의 일 실시예에 따른 n번째 프레임에서의 디스패리티 추정 방법을 나타낸 그래프이다.

[0068] 도 2를 참조하면, 디스패리티 추정 장치는 ROI 입력부(210), 탐색 윈도우 결정부(220) 및 디스패리티 근사화 유닛(225)을 포함할 수 있고, 디스패리티 근사화 유닛(225)은 디스패리티 추정부(230) 및 적응적 탐색 범위 결정부(240)를 포함할 수 있다.

[0069] ROI 입력부(210)는 사용자로부터 관심 영역(ROI)을 입력받는다. 관심 영역은 예를 들어 영상에 포함되는 특정 물체(예를 들어 얼굴 영역 등)로 지정될 수 있다(도 5a 참조).

[0070] 탐색 윈도우 결정부(220)는 관심 영역의 좌표 정보를 이용하여 디스패리티(disparity) 추정을 위해 사용되는 탐색 윈도우(Ω_n)를 결정한다. 탐색 윈도우는 관심 영역으로 지정된 물체를 검출하는 응용예에 적절한 알고리즘에 따라 그 모양이 결정될 수 있으며, 예를 들어 탐색 윈도우는 사각형 형상으로 결정될 수 있다. 또한 탐색 윈도우는 검출된 관심 영역의 모양과 무관하게 그 크기에만 영향을 받는 고정된 형태로 결정될 수도 있다.

[0071] 디스패리티를 연산하기 위해 사각형 형상의 탐색 윈도우는 하기 수학적 식 8을 이용하여 구해질 수 있다.

수학적 식 8

$$\Omega_n = \left\{ (i, j) \mid x_n^1 \leq i \leq \left\lfloor \frac{x_n^1 + x_n^2}{2} \right\rfloor, \left\lfloor \frac{y_n^1 + y_n^2}{2} \right\rfloor - \frac{H}{2} \leq j \leq \left\lfloor \frac{y_n^1 + y_n^2}{2} \right\rfloor + \frac{H}{2} \right\}$$

[0072]

[0073] 여기서, i와 j는 각각 x좌표와 y좌표를 나타내고, (x_n^1, y_n^1) 및 (x_n^2, y_n^2) 는 각각 n번째 프레임의 관심 영역 내에 위치한 좌측 상단 모서리의 좌표와 우측 하단 모서리의 좌표를 나타낸다. H는 짝수로 정의된 윈도우의 높이를 말하며, $\lfloor x \rfloor$ 는 x를 넘지 않는 최소 정수를 나타내고, A|B는 B 조건을 만족하는 A를 나타낸다.

[0074] 전술한 수학적 식 8을 이용하여 관심 영역으로부터 탐색 윈도우가 결정되면 디스패리티를 추정하기 위해 디스패리티 추정부(230)의 처리가 개시된다. 만일 사각형이 아닌 다른 모양으로 관심 영역이 검출된 경우에는 최소경계 사각형(minimum boundary rectangle)을 사용하여 검출된 관심 영역의 모양을 사각형으로 바꾸어 전술한 수학적 식 8에 따른 탐색 윈도우가 결정될 수도 있을 것이다.

[0075] 디스패리티 근사화 유닛(225)의 디스패리티 추정부(230)는 1차 테일러 급수(First-order Taylor Series)를 이용하여 각 프레임에 대한 디스패리티를 추정하고 근사화한다.

[0076] 디스패리티 추정부(230)는 탐색 윈도우 결정부(220)에 의해 결정된 탐색 윈도우(Ω_n)와 동영상에서의 템포럴(temporal) 정보(예를 들어, 이전에 구한 디스패리티 값)를 이용하여 유사측도(similarity measure)로부터 최소 값을 만족하는 디스패리티 기준값(d_0)을 찾고, 이를 이용하여 초기 디스패리티 추정값($\hat{d}_0$)을 추정하게 된다. 여기서, 디스패리티 기준값(d_0)은 좌측 및 우측 영상인 2개의 영상을 이용하여 템포럴 정보 없이 산출될 수 있으며, 초기 디스패리티 추정값($\hat{d}_0$)은 하기 수학적 식 9에 의해 추정될 수 있다.

수학적 식 9

$$\hat{d}_0 = \underset{d_0}{\operatorname{argmin}} \sum_{(p,q) \in \Omega_n} |I_L(p,q) - I_R(p+d_0,q)|, \quad 0 \leq \hat{d}_0 \leq D_0$$

[0077]

[0078] 여기서, I_L 과 I_R 은 2차원 평면으로 투사된 좌측 및 우측 2차원 영상을 나타내며, (p, q)는 탐색 윈도우 내의 픽셀의 좌표를 나타낸다. D_0 는 초기 디스패리티 추정값($\hat{d}_0$)을 구하기 위해 사용되는 초기 탐색 범위이다. 탐색 범위 내에 디스패리티를 구하기 위해 널리 사용되고 있는 유사측도는 예를 들어 SAD(sum of absolute

differences)가 사용될 수 있다.

- [0079] 디스패리티 초기 추정값($\hat{d}_0$)이 구해지면, 그 이후의 프레임에 대한 프레임별 디스패리티 추정값($\hat{d}_n$)은 하기 수학적 식 10에 의해 추정될 수 있다.

수학적 식 10

[0080]
$$\hat{d}_n = \tilde{d}_n + \Delta\hat{d}_n$$

- [0081] 여기서, $\tilde{d}_n$ 은 n번째 디스패리티 근사값을 나타내고, $\Delta\hat{d}_n$ 은 n번째 프레임의 디스패리티 보상값을 나타낸다.

- [0082] 수학적 식 10에서 1차 테일러 급수의 첫번째 항인 n번째 디스패리티 근사값을 나타내는 $\tilde{d}_n$ 은 하기 수학적 식 11로 정의될 수 있다.

수학적 식 11

[0083]
$$\tilde{d}_n = \hat{d}_{n-1} + d'_{n-1}$$

[0083]
$$d'_n = \hat{d}_n - \hat{d}_{n-1}$$

- [0084] 수학적 식 11의 두번째 항 d'_{n-1} 은 n-1 번째 프레임에서 추정된 디스패리티와 n-2번째 프레임에서 추정된 디스패리티 간의 차이 값이며, 수학적 식 11에서 정의된 바와 같이 미분의 의미를 내포하고 있다.

- [0085] 전술한 수학적 식 10의 두번째 항인 n번째 프레임의 디스패리티 보상값($\Delta\hat{d}_n$)은 하기 수학적 식 12로 정의될 수 있다.

수학적 식 12

[0086]
$$\Delta\hat{d}_n = \underset{\Delta\hat{d}_n}{\operatorname{argmin}} \sum_{(p,q) \in \Omega_n} |I_L(p + \tilde{d}_n, q) - I_R(p + \tilde{d}_n + \Delta\hat{d}_n, q)|, \quad -D_n \leq \Delta\hat{d}_n \leq D_n, 1 \leq n < N$$

- [0087] 여기서, D_n 은 초기 프레임을 제외한 n번째 프레임의 디스패리티를 구하기 위한 SAD에 사용되는 탐색 범위로서 후술될 적응적 탐색 범위 결정부(240)에 의해 결정되고, N은 비디오의 총 프레임 수를 나타낸다. 도 3은 n번째 프레임의 디스패리티를 추정하는 개념을 도식화하여 제시하고 있다.

- [0088] 다시 도 2를 참조하면, 적응적 탐색 범위 결정부(240)는 각 프레임에서 디스패리티를 구하기 위한 탐색 범위를 각 프레임별로 적응적 결정한다.

- [0089] 앞서 초기 디스패리티 추정값($\hat{d}_0$)을 추정하기 위한 수학적 식 9에서 사용되는 탐색 범위는 $0 \leq \hat{d}_0 \leq D_0$ 이지만, n번째 프레임의 디스패리티 보상값($\Delta\hat{d}_n$)을 연산하기 위한 수학적 식 12에서 사용되는 탐색 범위는 $-D_n \leq \Delta\hat{d}_n \leq D_n$ 로 차이를 가진다.

[0090] 이는, 최초 프레임의 디스패리티 보상값 $\Delta \hat{d}_0$ 을 이용하여 $\hat{d}_0$ 를 구하기 위해서는 넓은 범위 D_0 만큼 양의 방향으로 탐색을 해야 하는 반면, D_n 를 사용하여 구하는 $\hat{d}_n$ 의 경우에는 이미 산출된 최초 프레임의 디스패리티 보상값($\Delta \hat{d}_0$) 및 수학적 11 등을 사용하여 디스패리티 근사값 $\tilde{d}_n$ 을 구하고, $\tilde{d}_n$ 근처에서 좁은 영역인 D_n 만큼 양과 양의 방향으로 탐색하는 것으로 충분하다.

[0091] 또한, n-k부터 n번째 프레임까지의 인접한 k개 디스패리티들의 차이에 대한 추세를 반영하는 D_n 은 하기 수학적 13으로 정의될 수 있다.

수학적 13

$$D_n = D^0 + \lfloor |\tilde{D}_n| \rfloor$$

$$\tilde{D}_n = \frac{1}{k} \sum_{i=n-k+1}^n d'_i = \overline{d'_{n-1}} + \frac{d'_n - d'_{n-k}}{k}$$

[0092]

[0093] 여기서, D^0 는 탐색 범위의 오프셋(offset)을 나타내고, $|x|$ 는 x의 절대값을 의미한다. $\lfloor |\tilde{D}_n| \rfloor$ 는 n번째 프레임의 탐색범위의 보정값을 나타내며, $\tilde{D}_n$ 은 수학적 13에 정의된 바와 같이 k개의 디스패리티 값들의 차이들에 대한 평균값으로 정의된다. 또한 $\tilde{D}_n$ 을 정의하는 첫번째 항인 $\overline{d'_{n-1}}$ 는 d'_{n-k+1} 부터 d'_n 까지 k개의 평균값을 나타내고, $\tilde{D}_n$ 을 정의하는 첫번째 항은 연산 시간의 감소를 위해 이동평균필터(moving average filtering)로 재정의될 수 있다.

[0094] 또한 디스패리티들의 차이에 대한 잡음을 완화(smoothing)시키기 위해 k개의 평균값이 오프셋과 더해져서 매 프레임마다 적응적인 탐색범위가 적용되어 더 정확한 n번째 프레임의 디스패리티 보상값($\Delta \hat{d}_n$)을 구하는데 사용된다.

[0095] 전술한 바와 같이, 본 실시예에 따른 디스패리티 추정 장치는 동영상에서의 템포럴(temporal) 정보를 이용하여 각 프레임별로 디스패리티를 정확하게 추정하기 위해서 1차 테일러 급수(first-order Taylor series) 근사와 이동 평균 필터(moving average filter)를 사용하였으며, 이전에 추정된 디스패리티 값들과 추정된 디스패리티의 차이값들의 평균값을 현재 프레임의 디스패리티를 근사화하고 적응적인 탐색범위를 결정하는 데에 효율적으로 사용하는 특징을 가진다.

[0096] 도 4는 본 발명의 일 실시예에 따른 디스패리티 추정 방법을 나타낸 순서도이다.

[0097] 도 4를 참조하면, 단계 410에서 사용자로부터 관심 영역(ROI)이 입력된다. 관심 영역은 예를 들어 영상에 포함되는 특정 물체(예를 들어 얼굴 영역 등)로 지정될 수 있다.

[0098] 단계 420에서 탐색 윈도우 결정부(220)는 관심 영역의 좌표 정보를 이용하여 디스패리티(disparity) 추정을 위해 사용되는 탐색 윈도우(Ω_n)를 결정한다. 탐색 윈도우는 예를 들어 그 크기에만 영향을 받는 사각형 형상으로 결정될 수 있다.

[0099] 단계 430에서 디스패리티 근사화 유닛(230)은 최초 프레임에 대해 넓은 탐색 범위로 초기 디스패리티 추정값

($\hat{d}_0$)을 구한 후, 후속하는 각 프레임에 대해 적응적으로 탐색 범위(D_n)를 결정하고 각 프레임에 대한 프레임별 디스패리티 추정값($\hat{d}_n$)을 개별적으로 구한다.

- [0100] 일반적으로, 정지 영상에 대해서 수렴점(convergence point)을 맞추기 위해서는 기준이 되는 좌측 2차원 영상의 한 점과 일치하는 우측 2차원 영상의 한 점을 찾기 위해 넓은 탐색 범위에 대해 유사측도(similarity measure)를 사용하여 디스패리티(disparity)를 추정한다. 동영상에 대해서도 모든 프레임의 관심 영역에 수렴점을 맞추기 위해서 매번 정지 영상에서 디스패리티를 추정한 것처럼 모든 프레임에 대한 동일한 처리가 이루어질 수 있다.
- [0101] 그러나 매 프레임마다 넓은 탐색 범위가 적용된다면 디스패리티가 잘못 추정될 가능성이 높고, 연산시간 측면에서 효율적이지 못한 문제점이 있다.
- [0102] 이러한 문제점을 해소하기 위해, 본 실시예에 따른 동영상에서의 적응적 탐색 범위 예측을 이용한 디스패리티 추정 방법 및 장치는 현재 프레임의 디스패리티를 추정하기 위해 동영상에서의 템포럴 정보(예를 들어, 이전에 구한 디스패리티 값)를 사용함으로써 처음에만 넓은 탐색 범위를 적용하여 초기 디스패리티 추정값을 추정한 후, 이후의 프레임에 대해서는 상대적으로 좁은 탐색범위를 적용하여 프레임별 디스패리티 추정값을 추정함으로써 각 프레임의 디스패리티를 보다 정확하게 추정할 수 있을 뿐 아니라, 연산시간도 현저히 감소시킬 수 있는 장점이 있다.
- [0103] 고정된 좁은 탐색 범위를 설정하는 경우에는 넓은 탐색 범위보다 상대적으로 매우 작은 범위를 가질 수 있지만, 이 경우에도 움직임이 커서 디스패리티의 변화가 가장 클 때를 고려하여 탐색 범위를 설정할 수 밖에 없으며, 이는 움직임이 거의 없어 디스패리티 차이가 없는 프레임에서도 움직임이 클 때를 고려한 고정된 탐색 범위가 적용될 가능성이 있다.
- [0104] 따라서, 본 실시예에서는 동영상에서의 템포럴 정보를 이용하여 디스패리티의 변화가 가장 작을 때를 고려한 오프셋(offset)을 설정하고, 최근에 추정된 디스패리티의 변화 추세를 반영하여 탐색 범위를 적응적으로 변화시키는 방안을 제시함으로써 보다 정확한 디스패리티 추정이 가능한 장점도 있다.
- [0105] 도 5a 내지 도 5f는 본 발명의 일 실시예에 따른 디스패리티 추정 장치의 효과를 설명하기 위한 도면이다.
- [0106] 본 발명자는 본 실시예에 따른 디스패리티 추정 장치 및 방법을 이용한 3차원 이미지 처리 방법과 종래기술에 따른 3차원 이미지 처리 방법을 비교하기 위해 스테레오(stereo)로 설치한 두 개의 단안 카메라(즉, 마이크로 웹캠 시네마)로 각각 촬영한 4개의 비디오(도 5f의 (a) 내지 (d) 참조) 및 스테레오 카메라인 범블비 카메라로 촬영한 1개의 비디오 (도 5f의 (e) 참조)를 이용하였다.
- [0107] 각 실험 영상은 모두 정렬(rectification)된 영상을 사용하였으며, 각각 400장, 600장, 600장, 400장, 499장의 영상으로 구성되고, 두개의 단안 카메라로 촬영한 4개의 영상의 이미지당 사이즈는 960x588, 범블비 카메라로 촬영한 영상의 이미지당 크기는 1347x374이다.
- [0108] 실험 영상 1 내지 3(즉, 도 5f의 (a) 내지 (c))은 고정된 위치에 설치된 스테레오 카메라로부터 움직이는 관심 영역을 촬영한 영상이고, 실험 영상 4(즉, 도 5f의 (d))는 고정된 관심 영역을 스테레오 카메라를 움직이면서 촬영한 영상이며, 실험 영상 5(즉, 도 5f의 (e))는 움직이는 ROI를 차로 주행하면서 촬영한 영상이다. 실험 영상 1은 천장에 줄로 매달린 한 물체를 관심 영역으로 지정하고 이 물체가 원을 그리며 빙빙 도는 것을 촬영한 것이고, 실험 영상 2와 3은 사람 얼굴이 관심 영역이며, 다양한 방향으로 빠르게 혹은 느리게 움직이는 것을 촬영한 것이다. 실험 영상 4는 스테레오 카메라를 움직이면서 고정된 위치에 놓여진 다양한 물체를 촬영한 것이고, 이 물체들 중 사용자가 하나의 물체(즉, 도시된 object A)를 관심 영역으로 선택하였다. 실험 영상 5는 스테레오 카메라를 운전자 시야와 일치하게 차량에 고정시키고 주행하면서 앞에서 주행하는 차량을 촬영한 것이다.
- [0109] 또한, 비교를 위해 적절한 파라미터를 선정하기 위한 성능 평가 방법으로 RMS(Root Mean Square, E_{RMS})를 사용하였고, E_{RMS} 는 하기 수식식 14로 정의될 수 있다.

수학식 14

$$E_{\text{RMS}} = \sqrt{\frac{1}{N} \sum_{n=0}^{N-1} (\hat{d}_n - d_n)^2}$$

[0110]

[0111] 여기서, N은 실험에 사용된 비디오의 총 프레임 수이고, $\hat{d}_n$ 은 프레임별 디스패리티 추정값이며, d_n 은 n번째 프레임의 검증자료(ground truth)로서 본 발명자가 직접 확인하며 값을 측정하였다.

[0112] 실험 영상 1과 4는 검출 알고리즘에 의해 직사각형 형태로 관심 영역이 추출되었고, 실험 영상 2와 3은 검출 알고리즘에 의해 정사각형 형태로 ROI가 추출되었기 때문에 전술한 수학식 8을 이용하여 탐색 윈도우를 결정할 수 있었다.

[0113] 도 5a에는 검출 알고리즘에 의해 검출된 얼굴 영역에 대한 관심 영역과 탐색 윈도우가 도시되어 있다. 얼굴 영역을 검출하기 위한 검출 알고리즘은 다양할 수 있으나, 일 예로서 P. Viola 및 M. Jones의 논문 "Robust real-time face detection,"에서 제시한 검출 알고리즘을 설명하면 우선, Haar-like 특징 패턴으로부터 복수의 약 분류기(weak classifier)가 생성되고 복수의 약 분류기들은 Adaboost 알고리즘에 의해 강 분류기를 이루며, 강 분류기 (strong classifier)로 구성된 캐스케이드 구조에 의해 얼굴인지 아닌지 구분된다.

[0114] 적절한 파라미터 D^0 와 k를 선택하기 위해서 빠른 움직임을 촬영한 실험 영상 2를 사용하였다. 실험 영상 2는 사람의 얼굴을 ROI로 설정하고, 사람이 카메라를 향해 가까이 왔다가 다시 멀어지는 두 번의 상황을 빠른 속도의 움직임으로 촬영한 영상이다. 사람이 카메라를 향해 달려오다가 서서히 속도를 낮추면서 결국 470번째 프레임에 가까워 질수록 멈추고, 잠시 카메라 앞에 멈춘 후 다시 뒷걸음으로 점점 빨리 카메라로부터 멀어지는 상황을 연출했다. 이 때, 추출된 디스패리티 값들의 차이값에 대한 절대값 크기는 움직이는 물체의 속도가 정지하기 위해 느려졌다가 다시 빨라지기 때문에 470번째 프레임에 가까워 질수록 점점 작아지고, 잠시 0(zero)이었다가 가다가 그 이후에는 점점 커지는 양상을 보인다.

[0115] 우선, n번째의 탐색 범위 D_n 이 고정된 값인 오프셋(D^0)만이 아닌 디스패리티의 추세에 적응적으로 변화되는 $\tilde{D}_n$ 에 영향을 받는 이유에 대해서 도 5b 및 표 1을 통해 실험적으로 설명된다.

[0116] 도 5b는 D_n 과 D^0 가 일치하는 경우, D^0 의 변화에 따른 430번째 프레임부터 490번째 프레임까지의 디스패리티 보상값($\Delta \hat{d}_n$)을 나타낸 도면으로, 도시된 타원 A와 B에서의 결과는 탐색 범위가 적응적으로 변화되어야 하는 이유를 설명하고 있다.

[0117] 즉, 도 5b의 타원 A를 보면 $D^0=1$ 일 때 $\Delta \hat{d}_{447}$ 은 1로, $D^0=5$ 일 때 $\Delta \hat{d}_{447}$ 은 4로 보상되는 것을 확인할 수 있다. $D^0=1$ 일 때는 447번째 프레임 근처에서 보상값들이 클리핑되었지만, $D^0=3$ 일 때는 447번째 프레임에서 근처에서 보상값들이 클리핑되지 않은 것 또한 볼 수 있다.

[0118] 도 5c를 참조하면, 좁은 탐색 범위로부터 발생한 클리핑으로부터 잘못된 보상값이 산출되어 디스패리티를 잘못 추정하게 만드는 실험 결과가 확인된다. 참고로, 도 5c는 447번째 프레임에 대한 D^0 의 변화에 따른 결과 영상에 대한 것으로, (a)는 검증자료(ground truth)를 이용한 결과이고, (b)는 $D^0=1$ 일 때의 $\Delta \hat{d}_{447}$ 를 적용한 결과를 도 5c의 (a)에서의 사각형의 위치만큼 각각 확대하여 나타낸 것이고, (c)는 $D^0=5$ 일 때의 $\Delta \hat{d}_{447}$ 를 적용한 결과를 도 5c의 (a)에서의 사각형의 위치만큼 각각 확대하여 나타낸 것이다.

[0119] 도 5c의 (c)에서는 클리핑되지 않아 얼굴 영역에서 좌, 우 영상이 제대로 겹쳐지는 것을 볼 수 있지만, 그림 5c의 (b)에서는 클리핑되어 얼굴 영역에서 좌, 우 영상이 제대로 겹쳐지지 않는 것을 확인할 수 있으며, 이는 도

5c의 (b)에서는 탐색범위가 매우 작아 잘못된 보상값이 산출되어 결국 디스패리티가 부적절하게 추정되었기 때문이다. 그러나, 이러한 문제점을 해소하기 위해, 탐색 범위가 매우 클 필요는 없으며, 도 5b의 타원 B에 도시된 바와 같이 $D^{\circ}=3$ 일 때 $\Delta \hat{d}_{476}$ 이 -1로, $D^{\circ}=5$ 일 때 $\Delta \hat{d}_{476}$ 이 -5로 추정된 것을 확인할 수 있다. 이는, 탐색 범위가 필요 이상으로 큰 경우 디스패리티가 잘못 추정될 확률이 커짐을 보여준다.

[0120] 도 5d는 476번째 프레임에 대한 오프셋(D°)의 변화에 따른 결과 영상에 대한 것으로, (a)는 검증자료(ground truth)를 이용한 결과이고, (b)는 (a)의 사각형 위치만큼 $D^{\circ}=3$ 일 때의 결과를 확대한 영상이고, (c)는 (a)의 사각형 위치만큼 $D^{\circ}=5$ 일 때의 결과를 확대한 영상이다.

[0121] 적당한 탐색 범위가 적용된 도 5d의 (b)에서는 얼굴 영역에서 좌, 우 영상이 겹쳐지는 것을 볼 수 있지만, 도 5d의 (c)에서는 넓은 탐색 범위로 인해 발생한 에러 때문에 얼굴 영역에서 좌, 우 영상이 제대로 겹쳐지지 않음을 확인할 수 있다.

[0122] 하기 표 1은 탐색 범위의 오프셋에 따른 E_{RMS} 에 대한 결과를 비교한 것이다.

표 1

[0123]

	D°					
	1	2	3	4	...	10
E_{RMS}	62.92	1.64	1.41	1.54	...	1.67

[0124] 위 표1에서 보이는 바와 같이, $D^{\circ}=3$ 에서 가장 낮은 E_{RMS} 값을 나타내는데, 이는 $D^{\circ}=3$ 을 사용하여 디스패리티를 추정했을 때 가장 검증자료와 가깝다는 것을 의미한다. $D^{\circ}=5$ 일 때 클리핑 영역이 가장 나타나지 않지만 E_{RMS} 가 $D^{\circ}=3$ 일 때의 E_{RMS} 보다 낮지 않은 이유는 도 5b에서의 타원 B와 같이 넓은 탐색 범위 때문에 발생하는 잘못 추정된 디스패리티때문이다.

[0125] 하지만 항상 잘못된 디스패리티를 추정하는 것은 아니며, 도 5b에서 474번째 프레임 근처의 타원 C와 같이 일부 영역에서는 $D^{\circ}=3$ 일 때 보다 검증자료와 가까운 디스패리티를 추정하기도 한다. $D^{\circ}=1$ 일 때 큰 값의 E_{RMS} 를 나타내는 이유는 앞 프레임에서부터 매우 좁은 탐색 범위로부터 제대로 디스패리티가 추정되지 않고, 그 이후에도 계속 잘못된 디스패리티만을 추정하였기 때문이다.

[0126] 도 5d와 표 1을 참조하여 앞서 설명한 사항들을 고려할 때, 추정된 인접한 디스패리티들의 차이가 큰 구간에서는 탐색 범위를 넓힐 필요가 있고, 추정된 인접한 디스패리티들의 차이가 크지 않은 구간에서는 탐색 범위를 좁힐 필요성이 야기된다.

[0127] 도 5e는 인접한 디스패리티들의 개수 k 의 변화에 따른 430번째 프레임에서 490번째 프레임의 탐색범위의 보정값 $\lfloor |\tilde{D}_n| \rfloor$ 을 나타낸다.

[0128] 추정된 디스패리티들의 차이에 대한 k 개의 평균값들로 이루어진 $\lfloor |\tilde{D}_n| \rfloor$ 이기 때문에 $k=1$ 일 때의 결과에서 변동(fluctuation)이라 여겨지는 추정된 디스패리티들의 차이에 대한 튀는 값들이 그대로 발생하지만 $k=3$ 일 때의 결과에서는 $k=1$ 일 때의 결과에서 나타났던 큰 변동이 많은 정도 완화됨을 확인할 수 있다.

[0129] 다만, 전반적인 그래프에서 결과에 큰 영향을 주지 않는 무시할 수 있는 2프레임 정도의 지연이 발생되고 있다. $k=10$ 일 때의 결과에서는 $k=3$ 일 때의 결과보다 더 완화되어 잡음이라고 여겨지는 부분이 모두 없어진 것을 볼 수 있으나, 결과에 큰 영향을 주는 무시할 수 없는 5프레임 정도의 지연이 발생됨을 확인할 수 있다.

[0130] 하기 표 2는 오프셋을 3으로 고정하고 $n-k$ 부터 n 번째 프레임까지의 인접한 k 개 디스패리티들의 차이에 따른 RMS에 대한 결과를 비교한 것이다.

표 2

[0131]

	k					
	1	2	3	4	...	10
E_{RMS}	1.41	1.38	1.32	1.37	...	2.15

[0132]

표 2를 참조하면, k가 3에 가까워질수록 E_{RMS} 의 값이 점점 작아지다가 3보다 커질 경우 그래프의 지연 때문에 적응적인 탐색 범위의 크기가 잘못 적용되어 E_{RMS} 가 커지는 것을 볼 수 있다.

[0133]

도 5f는 예시된 다섯가지 실험 영상에 대해 좌측부터 차례대로 Chen의 방법(W.-C. Chen, F.-J. Hsiao, and C.-W. Lin, "An automatic parallax adjustment method for stereoscopic augmented reality systems," in Proc. Int. Symposium on 2010 9th Mixed and Augmented Reality, pp. 215-216, Seoul, Korea, Oct. 2010.), Xu의 방법(D. Xu, L. E. Corea, and P. Nasiopoulos, "Quality of experience for the horizontal pixel parallax," in Proc. 2012 Digest of Technical Papers Int. Conf. Consumer Electronics, pp. 394-395, Las Vegas, NV, Jan. 2012.) 및 본 실시예에 따른 방법에 대한 결과 영상을 나타낸다. 또한, 표 3은 도 5f에서의 결과에 적용된 추정된 디스패리티 값을 나타내고, 괄호 속의 값은 제안하는 방법에서 추정된 디스패리티 값과의 차이를 나타낸다.

표 3

[0134]

실험 영상	Chen의 방법	Xu의 방법	본 실시예
(a)	44(-42)	101(15)	86
(b)	8(-77)	120(35)	85
(c)	5(-33)	122(55)	38
(d)	50(-17)	70(18)	67
(e)	2(-5)	60(53)	52

[0135]

도 5f의 (a)를 참조하면, 본 실시예의 경우 추정된 디스패리티 86만큼 이동한 결과 관심 영역에 수렴점이 맞춰짐을 확인할 수 있으나, Chen이나 Xu 등의 방법을 적용하는 경우 각각 44 또는 101이라는 디스패리티가 적용되어 본 실시예와 비교할 때 각각 -42 또는 15라는 큰 음수차 또는 작은 양수차를 가지기 때문에 좌우 영상이 제대로 겹쳐지지 않음을 확인할 수 있다.

[0136]

또한 이러한 결과가 실험 영상 (b) 내지 (e)에서 공통됨을 확인할 수 있다. 이는 본 실시예의 경우 Chen이나 Xu의 방법과 달리 매 프레임마다 관심 영역의 디스패리티를 개별적으로 추정하여 그 프레임마다 추정된 디스패리티를 적용하기 때문이다. 다만 실험 영상 (d)의 경우 영상에 세개의 물체 A, B 및 C가 존재하며, 수렴점은 물체 A에 맞춰진 상태이며, 물체 B와 C는 각각 양수차 또는 음수차를 가지는 물체이므로 3차원 공간상에서 디스플레이 장치보다 앞 또는 뒤에 위치함을 확인할 수 있다.

[0137]

이와 같이, 본 실시예는 매 프레임마다 관심 영역의 디스패리티를 개별적으로 추정하여 그 프레임마다 추정된 디스패리티를 적용함으로써 좌우 영상의 관심 영역이 수렴점(즉, 매 프레임에서 디스패리티가 0(zero))에 맞춰질 수 있으며, 이를 통해 다른 종래의 방법들에 비해 눈의 피로도(visual fatigue)를 감소시킬 수 있고 눈의 편안함(visual comfort)을 증진시킬 수 있으며, 또한 주위 물체들의 상대적인 디스패리티 변화 때문에 3차원 효과도 더욱 느낄 수 있도록 하는 장점이 있다.

[0138]

상술한 디스패리티 추정 방법은 디지털 처리 장치에 내장되거나 설치된 프로그램 등에 의해 시계열적 순서에 따른 자동화된 절차로 수행될 수도 있음은 당연하다. 상기 프로그램을 구성하는 코드들 및 코드 세그먼트들은 당해 분야의 컴퓨터 프로그래머에 의하여 용이하게 추론될 수 있다. 또한, 상기 프로그램은 디지털 처리 장치가 읽을 수 있는 정보저장매체(computer readable media)에 저장되고, 디지털 처리 장치에 의하여 읽혀지고 실행됨으로써 상기 방법을 구현한다. 상기 정보저장매체는 자기 기록매체, 광 기록매체 및 캐리어 웨이브 매체를 포함

한다.

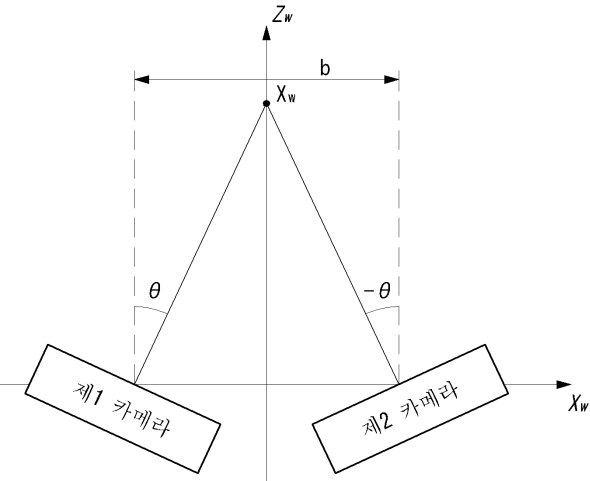
[0139] 상기에서는 본 발명의 실시예를 참조하여 설명하였지만, 해당 기술 분야에서 통상의 지식을 가진 자라면 하기의 특허 청구의 범위에 기재된 본 발명의 사상 및 영역으로부터 벗어나지 않는 범위 내에서 본 발명을 다양하게 수정 및 변경시킬 수 있음을 이해할 수 있을 것이다.

부호의 설명

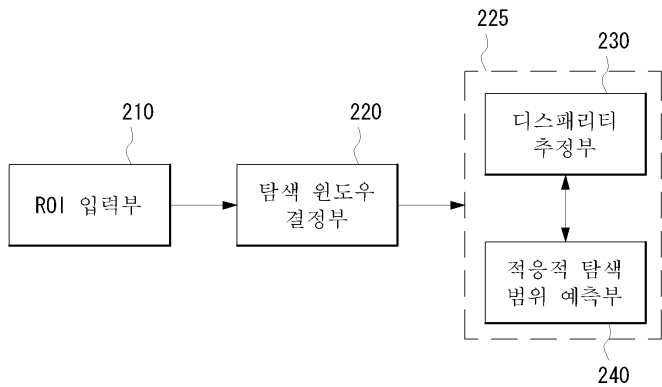
- [0140] 210 : ROI 입력부
220 : 탐색 윈도우 결정부
225 : 디스패리티 근사화 유닛
230 : 디스패리티 추정부
240 : 적응적 탐색 범위 결정부

도면

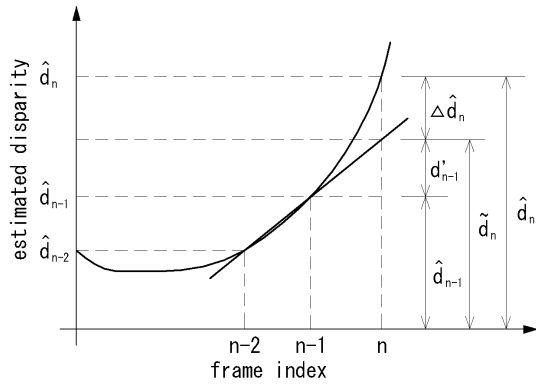
도면1



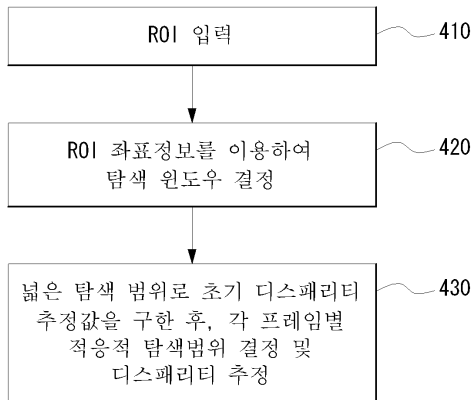
도면2



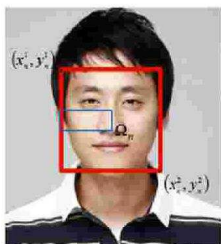
도면3



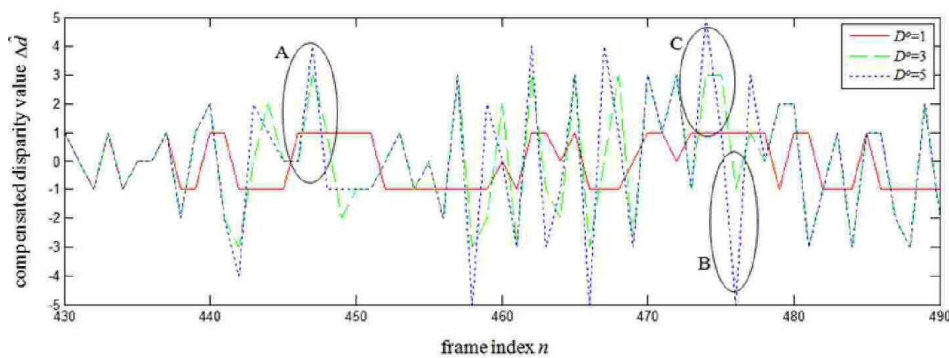
도면4



도면5a



도면5b



도면5c



(a)



(b)



(c)

도면5d



(a)

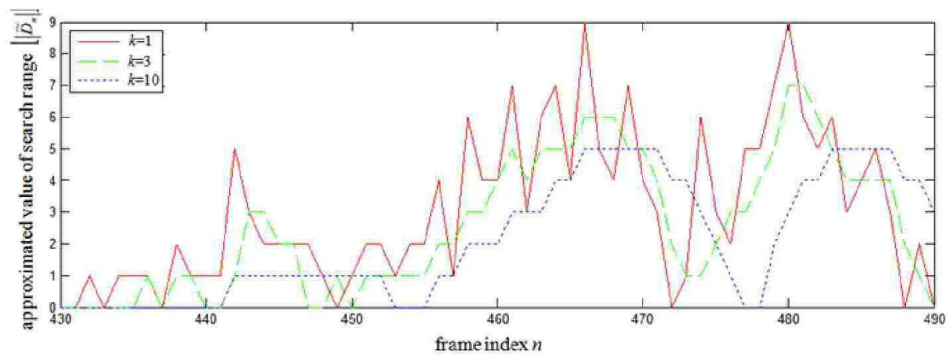


(b)



(c)

도면5e



도면5f

