

(19)대한민국특허청(KR)
(12) 공개특허공보(A)

(51) Int. Cl. (11) 공개번호 10-2006-0056050
H04N 5/262 (2006.01) (43) 공개일자 2006년05월24일

(21) 출원번호 10-2004-0095309
(22) 출원일자 2004년11월19일

(71) 출원인 주식회사 일리시스
서울특별시 강남구 삼성동 159-6 한국도심공항터미널빌딩 511호
(72) 발명자 장정훈
서울 서초구 방배2동 2761-2
(74) 대리인 송만옥

심사청구 : 있음

(54) 자동화된 360°파노라마 이미지 생성 방법

요약

본 발명은, 카메라를 한 곳에 고정시킨 채 물체를 회전시키면서 찍은 다수의 사진들을 이어 붙여서 360°를 포함하는 광각의 파노라마 이미지를 생성해내는 방법이다. 생성된 파노라마 이미지는 파노라마믹 VR 뷰어를 통해 보게 되며, 사용자에게 마치 촬영 장소에 서서 주위를 살펴보는 것과 같은 효과를 주게 된다. 본 발명은 아래와 같은 중요한 특징을 가지고 있다.

- 1) 로테이터와 같은 특수 장비 없이 삼각대만을 이용해서 고품질의 360° 파노라마 이미지를 생성할 수 있다.
- 2) 렌즈의 초점 거리에 대한 정보를 필요로 하지 않는다.
- 3) 광각 렌즈를 사용할 때 나타나는 이미지 왜곡 현상을 자동으로 보정한다.
- 4) 카메라가 지면과 수평을 이루지 않은 상태로 촬영된 사진들도 붙이는 것이 가능하다.
- 5) 사진 붙이기가 자동으로 이루어진다.
- 6) 입력 사진들 사이에 존재하는 밝기의 차이를 자동으로 보정한다.
- 7) 원통(Cylinder) 및 구(Sphere) 투영 파노라마 이미지를 생성한다.

대표도

도 1

색인어

템플릿 매칭, 360°파노라마 이미지, 호모그래피, LM, 8-파라메타, 투영.

명세서

도면의 간단한 설명

도 1은 본 발명의 실시예에 따른 360° 파노라마 이미지 생성 방법을 도식화한 플로우 차트.

도 2는 파라메타를 산출하는 과정을 보다 구체적으로 도식화한 플로우 차트.

도 3은 이미지 사이의 상대적인 카메라 회전각들을 산출하는 과정을 보다 구체적으로 도식화한 플로우 차트.

도 4는 360° 파노라마 이미지를 생성하는 과정을 구체적으로 도식화한 플로우 차트.

발명의 상세한 설명

발명의 목적

발명이 속하는 기술 및 그 분야의 종래기술

본 발명은 이미지의 생성 방법에 관한 것으로 특히, 자동화된 360° 파노라마 이미지 생성 방법에 관한 것이다.

파노라믹(Panoramic) VR(Virtual Reality; 가상 공간)을 구축하려면 크게 두 단계를 거쳐야 한다. 첫 번째는 파노라마 이미지를 얻는 단계이고, 두 번째는 파노라믹 VR 뷰어(Viewer)를 통해 전시하는 단계이다. 파노라믹 VR 뷰어는 입력으로 파노라마 이미지를 받고, 사용자가 특정한 시선 방향을 지정해주면, 그 시선 방향으로 보았을 때의 모습을 파노라마 이미지로부터 실시간으로 생성하여 보여준다.

최근에 인터넷(Internet) 상에서 파노라믹 VR이 보편화됨에 따라 많은 파노라믹 VR 뷰어가 출연하였다. 현재 많이 사용되고 있는 뷰어들은 Apple - QuickTime Player, Macromedia - Shockwave Player, PTVviewer, Anything3D - Viewer, iSeeMedia - Viewer, Zoomify - Viewer, iPIX - Viewer, MG System - Viewer 등이 있다.

이러한 두 단계 중 중요한 쪽은 파노라마 이미지를 얻는 과정이다. 카메라의 위치는 고정시킨 채, 물체를 회전시켜 가면서 연속적으로 얻은 사진들로부터 파노라마 이미지를 생성하려면 다음의 세 부분이 해결되어야 한다.

첫째, 이웃하는 사진들 사이의 픽셀의 대응 관계를 알아야 한다. (Image Registration)

둘째, 사진들을 원통 또는 구의 적당한 위치에 투영해야 한다. (Cylindrical / Spherical Projection)

셋째, 사진들을 원통 또는 구에 투영할 때, 사진들끼리 겹치는 부분이 부드럽게 이어지도록 해야 한다. (Image Blending)

초창기에 사용된 방법에서는 사진 촬영 시, 카메라가 지면과 완전히 수평을 이룬 상태에서 회전을 하고(Pure Panning), 카메라 렌즈의 초점 거리는 알고 있다는 조건하에서만 성립하는 알고리즘을 이용하였다.

렌즈의 초점 거리를 알고 있으면 원 사진을 원통에 투영할 수 있다. 카메라가 지면과 완전히 수평을 이룬 상태에서 회전을 한 경우, 이웃하는 원통 투영 이미지들 사이의 대응 관계는 단지 트랜스레이션(Translation)만으로 정의가 된다. 따라서, 트랜스레이션 양만 알면 이웃하는 원통 투영 이미지들끼리 이어 붙여서 파노라마 이미지를 만드는 것이 가능하게 된다. 이 이미지들 사이의 트랜스레이션 양은 템플릿 매칭(Template Matching)이라는 방법을 통해 쉽게 획득할 수 있다.

하지만, 이러한 방법을 적용하는 데에는 여러 가지 제약 사항들이 따른다. 우선, 렌즈의 초점 거리를 정확하게 알아야 한다. 그리고, 카메라가 지면과 완전히 수평을 이룬 상태에서 회전을 한 경우만을 허용하므로 로테이터(Rotator) 같은 특수 장비가 없으면 촬영하기가 어렵다.

또한, 카메라의 틸팅(Tilting)은 허용하지 않으므로 단일 열(Single Row)로 촬영된 파노라마 이미지만을 얻을 수 있다. 이러한 문제점들을 극복하기 위하여 렌즈의 초점 거리를 모르고, 카메라의 회전(Panning) 뿐만 아니라 틸팅이 존재하더라도 사진들을 붙일 수 있는 방법들을 개발하게 되었다.

카메라의 위치는 고정된 채, 회전만을 통해서 얻은 두 이미지 사이에는 3×3 의 행렬에 의해 정의되는 평면 투영 변환(Planar Projective Transformation) 관계가 있다. 투영 변환(Projective Transformation)은 스케일에 상관없이 정의되므로, 투영 변환 행렬은 8개의 파라메타로 구성되는데, 이것을 8-파라메타 호모그래피(8-Parameter Homography)라고 부른다.

지금까지 개발된 호모그래피를 구하는 방법은 다음과 같이 크게 두 부류로 나뉘어 진다.

그 첫 번째가 픽셀의 강도를 기반으로 한 방식(Pixel Intensity Based Method)으로, 이미지 사이의 겹치는 부분에서 픽셀들의 밝기 차이의 평균값이 최소가 되는 호모그래피(Homography)를 비선형 최소화(Nonlinear Minimization) 방법을 이용하여 구한다. 이 방식은 겹치는 영역의 픽셀들을 모두 이용하므로 텍스처(Texture)나 특징점(Feature Point)들이 적더라도 잘 동작하는 장점이 있으나, 계산 량이 많다는 단점이 있다.

두 번째가 특징을 기반으로 한 방식(Feature Based Method)으로, 이론적으로 8-파라메타 호모그래피는 4개 이상의 이미지 포인트 대응 쌍이 주어지면 구할 수 있다. 이미지들로부터 코너 포인트와 같은 특징 점들을 추출하고, 이미지 사이의 특징 점들의 대응 관계를 구한 후, 이것을 이용하여 호모그래피를 구한다. 이 방식은 속도가 빠르다는 장점은 있으나, 특징점들이 적을 경우에는 정확도가 떨어진다.

이미지 사이의 호모그래피를 구하였으면, 그것으로부터 초점거리와 카메라의 상대적인 회전각을 구하는 것이 가능하다. 이것을 카메라 자동 교정(Camera Auto-Calibration)이라고 한다. 초점거리와 카메라의 회전각을 알면, 원통 또는 구의 적당한 위치에 사진들을 투영하는 것이 가능하게 된다.

단일 열이 아닌 다중 열(Multi-Row)로 촬영된 사진들로부터 파노라마 이미지를 생성하는 경우, 2장 이상의 이미지들이 겹치는 영역이 발생하게 된다. 이런 경우 이미지 레지스트레이션 에러(Image Registration Error)가 누적되어 제대로 된 결과물을 얻기 힘들어지는데, 이러한 누적 에러를 줄이기 위해 글로벌 레지스트레이션(Global Registration) 방법이 개발되었다. 360° 파노라마 이미지를 생성하는 경우에도 유사한 문제가 발생한다. 이웃하는 이미지들 사이에는 비교적 정확하게 레지스트레이션(Registration)을 했더라도 에러가 누적되면 첫 번째 이미지와 마지막 이미지가 정확하게 레지스트레이션되지 않는다. 이런 경우에도 글로벌 레지스트레이션 방법을 이용해 해결하기도 한다.

발명이 이루고자 하는 기술적 과제

본 발명은 상기한 종래기술의 문제점을 해결하기 위해 제안된 것으로서, 로테이터와 같은 특수 장비 없이 삼각대만을 이용하여 고품질의 360° 파노라마 이미지를 생성할 수 있고, 렌즈의 초점 거리에 대한 정보를 필요로 하지 않으며, 광각 렌즈를 사용할 때 나타나는 이미지 왜곡 현상을 자동으로 보정할 수 있으며, 카메라가 지면과 수평을 이루지 않은 상태로 촬영된 사진들도 붙이는 것이 가능하며, 입력 사진들 사이에 존재하는 밝기의 차이를 자동으로 보정하며, 원통(Cylinder) 및 구(Sphere) 투영 파노라마 이미지를 생성할 수 있는 자동화된 360° 파노라마 이미지 생성 방법을 제공하는 것을 그 목적으로 한다.

발명의 구성 및 작용

상기의 목적을 달성하기 위해 본 발명은, 360° 방향으로 복수의 사진을 촬영하는 단계; 촬영된 상기 복수의 사진으로부터 선택된 임의의 이웃하는 사진 한 쌍을 이용하여 카메라의 초점거리와 왜곡률 및 두 이미지 사이의 회전각으로 이루어진 파라메타를 산출하는 단계; 상기 산출된 파라메타를 이용하여 각각의 이웃하는 사진 쌍으로부터 이미지 사이의 상대적인 카메라 회전각들을 산출하는 단계; 상기 이미지 사이의 상대적인 회전각들을 이용하여 각 이미지에 대응하는 카메라 파라메타를 계산하는 단계; 상기 카메라 파라메타를 이용하여 이웃하는 이미지 사이의 밝기 차이를 보정하는 단계; 밝기 차이가 보정된 각 이미지를 단일 원통에 투영하는 단계; 상기 투영 결과, 겹치는 영역에서 이미지 블랜딩을 실시하여 미완성 파노라마 이미지를 형성하는 단계; 사진 열의 첫 번째 이미지를 이용하여 원통 투영 이미지를 생성하는 단계; 및 상기 원통 투영 이미지와 상기 미완성 파노라마 이미지를 이용하여 360° 파노라마 이미지를 생성하는 단계를 포함하는 360° 파노라마 이미지 생성 방법을 제공한다.

본 발명은, 렌즈 왜곡 현상을 고려하고, 카메라 틸트를 고려하며, 페어 와이어 국보적인 방식을 이용하여 이미지 레지스트레이션을 수행한다. 또한, 입력 사진들 사이에 존재하는 밝기 차이를 자동으로 보정하며, 360° 파노라마 이미지를 생성한다.

이하, 본 발명이 속하는 기술분야에서 통상의 지식을 가진 자가 본 발명을 보다 용이하게 실시할 수 있도록 하기 위하여 본 발명의 바람직한 실시예를 첨부한 도면을 참조하여 상세하게 설명한다.

도 1은 본 발명의 실시예에 따른 360° 파노라마 이미지 생성 방법을 도식화한 플로우 차트로서, 이를 참조하여 파노라마 이미지 생성 과정을 살펴본다.

먼저, 파노라마 이미지를 생성하기 위해 360° 방향으로 복수의 사진을 촬영한다(S101). 촬영된 사진들로부터 임의의 이웃하는 사진 한 쌍을 선택한 후, 선택된 임의의 이웃하는 사진 한 쌍을 이용하여 파라메타를 산출한다(S102). 파라메타는 카메라의 초점거리와 왜곡률 및 두 이미지 사이의 회전각을 포함한다.

산출된 파라메타를 이용하여 각각의 이웃하는 사진 쌍으로부터 이미지 사이의 상대적인 카메라 회전각들을 산출한다(S103). 이미지 사이의 상대적인 회전각들을 이용하여 각 이미지에 대응하는 카메라 파라메타를 계산한다(S104).

카메라 파라메타를 이용하여 이웃하는 이미지 사이의 밝기 차이를 보정한다(S105). 이어서, 밝기 차이가 보정된 각 이미지를 단일 원통에 투영한다(S106).

상기 투영 결과, 겹치는 영역에서 이미지 블렌딩을 실시하여 미완성 파노라마 이미지를 형성한다(S107). 사진 열의 첫 번째 이미지를 이용하여 원통 투영 이미지를 생성한 후(S108), 원통 투영 이미지와 미완성 파노라마 이미지를 이용하여 360° 파노라마 이미지를 생성한다(S109).

도 2는 파라메타를 산출하는 과정을 보다 구체적으로 도식화한 플로우 차트이다.

도 2를 참조하면, 파라메타를 산출하는 단계는, 임의의 이웃하는 사진 한 쌍에 대해 템플릿 매칭(Template matching)을 통한 두 이미지 사이의 트랜스레이션(Translation) 양을 계산하는 단계(S201)와, LM(Levenberg-Marquardt) 최소화를 통한 두 이미지 사이의 회전각을 계산하는 단계(S202)와, LM 최소화를 통한 렌즈의 왜곡률 및 두 이미지 사이의 8-파라메타 호모그래피(Homography)를 계산하는 단계(S203)와, 카메라 자동 교정(Camera auto-calibration)을 통한 카메라의 초점거리를 계산하는 단계(S204)로 이루어진다.

도 3은 이미지 사이의 상대적인 카메라 회전각들을 산출하는 과정을 보다 구체적으로 도식화한 플로우 차트이다.

도 2를 참조하면, 이미지 사이의 상대적인 카메라 회전각들을 산출하는 단계는, 산출된 파라메타를 이용하여 각각의 이웃하는 사진 쌍으로부터 렌즈 왜곡이 보정된 이미지를 생성하는 단계(S301)와, 렌즈 왜곡이 보정된 이미지를 이용하여 원통 투영 이미지를 생성하는 단계(S302)와, 이미지 회전 후, 템플릿 매칭을 통한 두 이미지 사이의 트랜스레이션 양을 계산하는 단계(S303)와, LM 최소화를 통한 두 이미지 사이의 3-파라메타 호모그래피를 계산하는 단계(S304)로 이루어진다.

도 4는 360° 파노라마 이미지를 생성하는 과정을 구체적으로 도식화한 플로우 차트이다.

도 4를 참조하면, 360° 파노라마 이미지를 생성하는 단계는, 원통 투영 이미지와 미완성 파노라마 이미지를 이용하여 템플릿 매칭을 통한 두 이미지 사이의 트랜스레이션 양을 계산하는 단계(S401)와, 이미지 합체 후 절단하는 단계(S402)로 이루어진다.

한편, 템플릿 매칭 및 LM 최소화의 수행 시 계산량을 줄이기 위해 이미지 피라미드를 이용한다.

이하에서는 상기한 본 발명의 360°파노라마 이미지 생성 과정을 수학적식을 이용하여 구체적으로 살펴본다.

렌즈에 의한 왜곡이 없는 일반적인 카메라에 의하여 3차원 공간상의 점 $\mathbf{X}=(X,Y,Z)$ 이 2차원의 이미지 면에 투영되어 얻어진 점을 $\mathbf{x}=(x,y)$ 라고 하면, \mathbf{X} 와 \mathbf{x} 와의 관계는 하기의 수학적식1 및 수학적식2와 같이 모델링될 수 있다.

$$\text{수학적식 1}$$

$$\mathbf{u} = \begin{bmatrix} u \\ v \\ w \end{bmatrix} = \mathbf{K}(\mathbf{R}\mathbf{X} + \mathbf{t})$$

$$= \begin{bmatrix} f & s & x_c \\ 0 & af & y_c \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

수학식 2

$$x = \frac{u}{v}, \quad y = \frac{v}{w}$$

여기에서 \mathbf{u} 는 \mathbf{x} 를 단일 좌표(Homogeneous Coordinates)로 표시한 벡터이고, \mathbf{K} 는 카메라 교정 행렬(Camera Calibration Matrix)로써, 카메라 좌표계 상에서 3D(3차원) 포인트와 그것이 투영되어 얻어진 이미지 면에서의 2D(2차원) 포인트와의 관계를 나타내고, \mathbf{R} 과 \mathbf{t} 는 카메라의 회전 행렬(Rotation Matrix)과 트랜스레이션 벡터(Translation Vector)로써, 월드 좌표계와 카메라 좌표계 사이의 변환 관계를 나타낸다. 행렬 \mathbf{K} 에 포함되어 있는 파라메타들을 카메라의 내부 파라메타(Internal Parameter)라고 부르는데, f 는 초점거리(Focal Length), a 는 픽셀의 종횡비(Aspect Ratio), s 는 스큐(Skew), (x_c, y_c) 는 주된 점(Principal Point)를 나타낸다.

이제 카메라의 투영 중심(Projection Center)를 고정시키고 카메라를 일정량만큼 회전시킨 후에 얻은 이미지 I_1 과 회전시키기 전에 얻은 이미지 I_0 사이의 픽셀의 대응 관계에 대해 알아본다.

회전하기 전의 카메라의 회전 행렬과 트랜스레이션을 각각 R_0 와 t_0 라고 하고, 회전한 후의 것을 R_1 , t_1 이라고 한다. 편의상 카메라의 투영 중심을 월드 좌표계의 원점과 일치시키면, $t_0=t_1=0$ 가 된다. 3차원 공간상의 한 점 X 가 I_0 와 I_1 에 투영되어 얻어진 점의 단일 벡터(Homogeneous Vector)를 각각 u_0 와 u_1 이라고 하면, 이들은 하기의 수학식3과 같은 관계를 갖는 것을 알 수 있다.

수학식 3

$$u_1 = KR_1X = KR_1R_0^{-1}K^{-1}u_0 = KR_{01}K^{-1}u_0 = Hu_0$$

즉, 카메라를 회전시켜서 얻은 두 이미지 사이의 픽셀의 대응 관계는 3×3 의 행렬 H 에 의해 표현되는데, 이것을 평면적인 호모그래피(Planar Homography)라고 한다. 단일 벡터에 0이 아닌 임의의 실수를 곱하더라도 그것에 대응되는 이미지 포인트는 같다는 점을 고려하면, H 또한 스케일에 상관없이 정의된다는 것을 알 수 있다. 따라서, 평면적인 호모그래피는 9개가 아닌 8개의 파라메타에 의해 정의된다. 실제로 $H=KR_{01}K^{-1}$ 에서 K 에 포함되어 있는 파라메타 수는 5개이고, R_{01} 에 포함되어 있는 파라메타 수는 3개(하기의 수학식4 참조)이므로, H 를 구성하는 독립적인 파라메타 수는 8개가 됨을 알 수 있다. 이러한 호모그래피를 8-파라메타 호모그래피라고 한다.

두 이미지가 주어졌을 때, 두 이미지 사이의 픽셀의 대응 관계 H 를 알아내는 것을 이미지 레지스트레이션이라고 한다. H 는 일반적으로 4개 이상의 대응점이 주어지면 구할 수 있으나, 정확한 H 를 구하기 위해서는 세심한 주의가 필요하다.

이제 좀 더 단순화 된 모델에 대해 알아본다. 판매되고 있는 대부분의 카메라는 다음의 조건을 만족한다.

- 1) 픽셀들은 대부분 정사각형 픽셀(Square Pixel)이므로 픽셀의 종횡비는 $a=1$ 을 만족한다.
- 2) 이미지 면을 이루는 CCD(Charge Coupled Device) 셀들은 대부분 직각으로 배치가 되므로 스큐 성분은 $s=0$ 를 만족한다.
- 3) 카메라의 광축(Optical Axis)은 CCD 어레이(Array)의 중심을 통과한다고 가정해도 무방하므로 주된 점 (x_c, y_c) 는 이미지의 중심과 일치한다고 본다.

따라서, 단순화된 모델에서는 H 에 포함된 독립적인 파라메타의 수는 4개(초점거리 f 와 R_{01} 에 포함된 3개의 회전 각)가 되는데, 이것을 4-파라메타 호모그래피라고 한다.

참고로 일반적인 회전 행렬 R은 하기의 수학식4와 같이 분해된다.

수학식 4

$$R = R(\psi)R(\theta)R(\phi)$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

여기에서 R(Ψ), R(Θ), R(Φ)는 각각 X, Y, Z축을 중심으로 한 회전 변환을 나타내는 행렬이며, Ψ, Θ, Φ를 율러 각(Euler Angle)이라고 한다.

단순화된 모델에서 평면 호모그래피(Planar Homography) H가 주어졌을 때, K와 R을 구하는 방법은 하기의 수학식들을 통해 표현할 수 있다. 행렬 K는 하기의 수학식5와 같이 같이 분해될 수 있다.

수학식 5

$$K = K_o K_f = \begin{bmatrix} 1 & 0 & x_o \\ 0 & 1 & y_o \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

수학식5를 수학식3에 대입하면 수학식6과 같이 된다.

수학식 6

$$K_f R K_f^{-1} = K_o^{-1} H K_o = H = \begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

따라서, R은 하기의 수학식7과 같이 표현된다.

수학식 7

$$R = \begin{bmatrix} h_{11} & h_{12} & h_{13}/f \\ h_{21} & h_{22} & h_{23}/f \\ h_{31}f & h_{32}f & h_{33} \end{bmatrix}$$

행렬 R은 직교 행렬(Orthogonal Matrix)이므로 하기의 수학식8을 만족한다.

수학식 8

$$h_{11}^2 + h_{12}^2 + \frac{h_{13}^2}{f^2} = h_{31}^2 f^2 + h_{32}^2 f^2 + h_{33}^2$$

수학식8로부터 f를 구한 후, 수학식7에 대입하여 R을 구할 수 있다.

한편, 실제로 H로부터 f와 R을 구하는데 있어서, 다음과 같은 주의가 필요하다.

- 1) H가 스케일에 상관없이 주어지므로, 수학식7로부터 R을 구한 후, $RR^T=I$ 가 되도록 R의 스케일 값을 조정해야 한다.
- 2) 주어진 H가 4-파라메타 호모그래피 모델을 만족하지 않는 경우에는 수학식7을 통해 구한 R은 정확하게 직교 행렬이 되지 않는다. 따라서, 일차적으로 구한 R로부터 SVD(Singular Value Decomposition)를 이용해 근사화된 직교 행렬을 구해야 한다.

이상에서 언급한 카메라 모델에는 렌즈에 의한 왜곡 현상은 고려되어 있지 않다. 현실적으로 360° 파노라마 영상을 얻기 위하여 촬영 시 광각 렌즈를 많이 사용하기 때문에, 렌즈 왜곡 현상을 고려해야만 보다 넓은 범위에서 알고리즘의 적용이

가능하다. (x_d, y_d) 를 원 이미지(즉, 렌즈 왜곡이 있는 이미지)에서의 픽셀의 좌표라고 하고, (x_u, y_u) 를 렌즈 왜곡이 제거된 이미지에서의 대응되는 픽셀 좌표라고 한다. 그리고, (x_c, y_c) 를 원 이미지의 중심 좌표라고 하면, (x_d, y_d) 이 주어졌을 때, (x_u, y_u) 는 하기의 수학식9와 같다.

수학식 9

$$x_u = (x_d - x_c)(1 + \kappa_1 r_d^2 + \kappa_2 r_d^4 + \dots) + x_c,$$

$$y_u = (y_d - y_c)(1 + \kappa_1 r_d^2 + \kappa_2 r_d^4 + \dots) + y_c,$$

$$r_d = \sqrt{(x_d - x_c)^2 + (y_d - y_c)^2}.$$

여기에서 K_1, K_2 를 방사 왜곡 계수(Radial Distortion Coefficients)라고 한다. 대부분의 경우 첫 번째 방사 왜곡 계수 항만 고려해도 충분하므로 $K=K_1$ 이라고 하면, 수학식9는 하기의 수학식10과 같이 간략화 된다.

수학식 10

$$x_u = (x_d - x_c)(1 + \kappa r_d^2) + x_c,$$

$$y_u = (y_d - y_c)(1 + \kappa r_d^2) + y_c.$$

반대로 (x_u, y_u) 이 주어졌을 때, (x_d, y_d) 는 하기의 수학식11과 같다.

수학식 11

$$x_d = \frac{x_u - x_c}{1 + \kappa r_d^2} + x_c, \quad y_d = \frac{y_u - y_c}{1 + \kappa r_d^2} + y_c.$$

여기에서, r_d 는 K 와 하기의 수학식12와 같으며,

수학식 12

$$r_d^2 = h(\kappa) + \frac{1}{9h(\kappa)\kappa^2} - \frac{2}{3\kappa},$$

$h(K)$ 와 하기의 수학식13과 같으므로,

수학식 13

$$h(\kappa) = \left(\sqrt{\left(\frac{1}{27\kappa^3} + \frac{r_u^2}{2\kappa^2} \right)^2 - \frac{1}{729\kappa^6}} + \frac{1}{27\kappa^3} + \frac{r_u^2}{2\kappa^2} \right)^{\frac{1}{3}},$$

r_u 는 (x_c, y_c) 및 (x_u, y_u) 와 하기의 수학식14와 같은 관계를 갖는다.

수학식 14

$$r_u^2 = (x_u - x_c)^2 + (y_u - y_c)^2$$

이하에서는 카메라 파라미터를 추출한다(Camera Parameter Extraction).

사용자가 촬영한 사진들이 입력으로 주어지면, 파노라마 이미지 생성 도구는 임의의 이웃하는 한 쌍의 사진들을 가지고 이미지 레지스트레이션을 수행하여 8-파라미터 호모그래피 H 와 방사 왜곡 계수 K 를 구한 후, H 로부터 수학식8을 이용하여 초점거리 f 를 구한다.

사진 촬영 시, 중간에 렌즈를 교환하거나 줌인(Zooming)을 하는 경우는 없다고 가정하면, f와 K값은 고정되기 때문에 임의의 한 쌍의 이미지에서 구한 f와 K를 모든 이미지에 공통적으로 적용해도 무방하다.

본 발명에서 사용하는 이미지 레지스트레이션 방법은 다음과 같다.

주어진 한 쌍의 이미지에서 첫 번째 것을 I_0 , 두 번째 I_1 것을 이라고 한다. 또한, I_0 에 속하는 임의의 픽셀 $\mathbf{x}=(x,y)$ 에 대응하는 I_1 의 픽셀의 좌표를 $\mathbf{x}^\wedge=(x^\wedge,y^\wedge)$ 라고 한다. 그리고, \mathbf{x} 를 \mathbf{x}^\wedge 으로 변환시키는 평면 호모그래피 H는 하기의 수학식15와 같이 표시된다.

수학식 15

$$H = \begin{bmatrix} m_1 & m_2 & m_3 \\ m_4 & m_5 & m_6 \\ m_7 & m_8 & 1 \end{bmatrix}$$

H가 스케일에 상관없이 정의되기 때문에, H의 마지막 요소를 1로 놓았다. 만약, 렌즈 왜곡이 없다고 가정하고 \mathbf{x} 의 단일 좌표계를 $(x,y,1)$ 로 놓으면, \mathbf{x} 와 \mathbf{x}^\wedge 의 대응 관계는 수학식3을 이용하여 하기의 수학식16으로 표현할 수 있다.

수학식 16

$$x = \frac{m_1x + m_2y + m_3}{m_7x + m_8y + 1}, \quad y = \frac{m_4x + m_5y + m_6}{m_7x + m_8y + 1}$$

H를 구하기 위하여, 하기의 수학식17과 같이 정의되는 비용 함수(Cost Function) e를 고려한다.

수학식 17

$$e(m_1, \dots, m_8) = \frac{1}{N_R} \sum_{\mathbf{x} \in R} e_{\mathbf{x}}^2$$

$$= \frac{1}{N_R} \sum_{\mathbf{x} \in R} (I_1(x_{\mathbf{x}}, y_{\mathbf{x}}) - I_0(x_{\mathbf{x}}, y_{\mathbf{x}}))^2$$

여기에서 R은 I_0 와 I_1 사이의 겹치는 영역을 나타내고, N_R 은 R에 속하는 픽셀의 수를 나타내며, $I(x,y)$ 는 이미지 I에서 (x,y) 의 위치에 존재하는 픽셀의 밝기 값을 나타낸다.

그러면, 비용 함수 e를 최소화하는 H를 구하고자 하는 최적의 H라고 본다. 3차원 공간상의 한 점이 서로 다른 이미지로 투영되어 얻어진 픽셀들은 모두 같은 밝기를 갖는다고 가정하면, 위에서 제시한 비용 함수 e를 최소화하는 H가 찾고자 하는 최적의 H가 되는 것은 타당하다.

적당한 H의 초기값이 주어졌을 때, 수치 해석적인 비선형 최소화 방법을 이용하여 e를 최소화하는 H를 구할 수 있는데, 대표적으로 Levenberg-Marquardt 최소화 방법이 있다.

참고로, Levenberg-Marquardt (이하 LM이라 함) 최소화 방법은 뉴턴 최소화(Newton Minimization) 방법을 약간 변형한 방법이다.

이하에서는, 뉴턴 최소화 방법을 간략하게 살펴 본다.

최소화하고자 하는 비용 함수 e가 하기의 수학식18과 같은 형태로 주어졌다고 가정한다.

수학식 18

$$e(\mathbf{m}) = ||\mathbf{e}(\mathbf{m})||^2 = \sum_{i=1}^N e_i^2(\mathbf{m})$$

여기에서 $m=(m_1, \dots, m_M)$ 은 파라메타 벡터(Parameter Vector)이다. 현재 주어진 파라메타 벡터의 m_0 값을 라고 할 때,

수학식 19

$$\|e(m_0 + \Delta m)\| < \|e(m_0)\|$$

수학식19를 만족하는 벡터 Δm 을 찾았다고 가정하면, m_0 를 $m_0 \leftarrow m_0 + \Delta m$ 으로 갱신한 후, 다시 수학식19를 만족하는 Δm 을 찾는 것을 반복하면, 결국 e 의 국부적인 최소값(Local Minimum)에 도달하게 될 것이다. Δm 을 찾는 여러 가지 방법이 있지만, 뉴턴 최소화 방법에서는 다음과 같은 방식으로 찾는다.

$e(m_0 + \Delta m)$ 을 m_0 주위에서 테일러 시리즈(Taylor Series)로 1차 항까지만 전개하면 하기의 수학식20과 같이 근사화 될 수 있다.

수학식 20

$$e(m_0 + \Delta m) \approx e(m_0) + J(m_0)\Delta m.$$

여기에서 행렬 $J(m)$ 을 야코비안 행렬(Jacobian Matrix)이라고 부르고, 하기의 수학식21과 같이 주어진다.

수학식 21

$$J(m) = \begin{bmatrix} \frac{\partial e_1}{\partial m_1} & \frac{\partial e_1}{\partial m_2} & \dots & \frac{\partial e_1}{\partial m_M} \\ \frac{\partial e_2}{\partial m_1} & \frac{\partial e_2}{\partial m_2} & \dots & \frac{\partial e_2}{\partial m_M} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial e_N}{\partial m_1} & \frac{\partial e_N}{\partial m_2} & \dots & \frac{\partial e_N}{\partial m_M} \end{bmatrix}.$$

뉴턴 최소화 방법에서는 $\|e(m_0) + J(m_0)\Delta m\|$ 을 최소화하는 Δm 을 찾는 바, Δm 은 하기의 수학식22와 같이 주어진다.

수학식 22

$$\Delta m = -J^+ e(m_0) = -(J^T J)^{-1} J^T e(m_0).$$

여기에서 $J=J(m_0)$ 이고, J^+ 를 J 의 의사 역(Pseudo-Inverse)이라고 한다.

수학식22에서 $A=J^T J$ 라고 하고, a_{ij} 를 행렬 A 의 (i,j) 에 위치하는 요소라고 하면, LM 최소화 방법에서는 행렬 A 를 하기의 수학식23과 같이 주어지는 행렬 \hat{A} 으로 교체한다.

수학식 23

$$\hat{A} = \begin{bmatrix} (1+\lambda)a_{11} & a_{12} & \dots & a_{1M} \\ a_{21} & (1+\lambda)a_{22} & \dots & a_{2M} \\ \vdots & \vdots & \ddots & \vdots \\ a_{M1} & a_{M2} & \dots & (1+\lambda)a_{MM} \end{bmatrix}.$$

λ 의 초기 값은 보통 10^{-3} 으로 설정한다. 만약, 현재 설정된 λ 값을 이용하여 Δm 을 구한 결과, e 의 값이 줄어들었다면 $\lambda \leftarrow 0.1\lambda$ 로 갱신하고, e 의 값이 늘어났다면 $\lambda \leftarrow 10\lambda$ 로 갱신한다. λ 값이 매우 작아지면, LM 최소화 방법은 뉴턴 최소화 방법과 유사하게 되고, λ 값이 매우 커지면, LM 최소화 방법은 구배에 맞는 최소화(Gradient Decent Minimization) 방법과 유사하게 됨을 관찰할 수 있다. 일반적으로, LM 최소화 방법은 뉴턴 최소화 방법보다 수렴 속도가 빠르며 안정적으로 동작하는 것으로 알려져 있다.

LM 최소화 방법을 이용하여 수학식17을 최소화하려면, 수학식21의 야코비안 행렬을 구해야 한다. J 의 (i,j) 에 위치하는 J_{ij} 라고 하면, 요소를 체인 룰(Chain rule)에 의하여 하기의 수학식24와 같이 주어진다.

수학식 24

$$J_{ij} = \frac{\partial e_i}{\partial m_j} = \frac{\partial I_1(x_i, y_i)}{\partial m_j} = \frac{\partial I_1}{\partial x_i} \frac{\partial x_i}{\partial m_j} + \frac{\partial I_1}{\partial y_i} \frac{\partial y_i}{\partial m_j}$$

여기에서 $(\frac{\partial I_1}{\partial x_i}, \frac{\partial I_1}{\partial y_i})$ 는 (x_i, y_i) 의 위치에서 I_1 의 구배(gradient)가 된다.

본 발명에서는 렌즈 왜곡을 고려한 이미지 레지스트레이션 방법을 제공하는 바, 렌즈 왜곡 현상을 고려하려면 수학식16과 수학식17은 각각 수학식25와 수학식26으로 바뀌어야 한다.

수학식 25

$$\hat{x}_u = \frac{m_1 x_u + m_2 y_u + m_3}{m_7 x_u + m_8 y_u + 1}, \quad \hat{y}_u = \frac{m_4 x_u + m_5 y_u + m_6}{m_7 x_u + m_8 y_u + 1}$$

수학식 26

$$e(m_1, \dots, m_8, \kappa) = \frac{1}{N_R} \sum_{(x_d, y_d) \in R} (I_1(x_d, y_d) - I_0(x_d, y_d))^2$$

여기에서 $(x_d, y_d) \leftrightarrow (x_u, y_u)$ 의 변환 관계(또는 $(x^{\wedge}_d, y^{\wedge}_d) \leftrightarrow (x^{\wedge}_u, y^{\wedge}_u)$)는 수학식10과 수학식11에서 주어진 것과 같다.

참고로, 수학식24에서 $\frac{\partial \hat{x}_i}{\partial m_j}$ (또는 $\frac{\partial \hat{y}_i}{\partial m_j}$)를 구하는데 있어서, 분석적인(Analytic) 방법에 의하여 $\frac{\partial \hat{x}_i}{\partial m_j}$ 의 일반적인 형태를 얻기 힘들면, 수식적인(Numerical) 방법에 의하여 구해야만 한다.

일반적으로 함수 $f(x)$ 가 주어졌을 때, $x=x_0$ 에서 df/dx 는 하기의 수학식27과 같이 근사화 될 수 있다.

수학식 27

$$\left. \frac{df}{dx} \right|_{x=x_0} \approx \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

여기에서 Δx 는 $|x_0|$ 에 비해 상대적으로 매우 작은 양의 실수로써, 하기의 수학식28과 같이 설정하면 좋은 결과를 얻을 수 있다고 알려져 있다.

수학식 28

$$\Delta x = \max(|10^{-4} x_0|, 10^{-6})$$

수학식17의 비용 함수 e 를 LM 최소화 방법에 의해 최소화하여 우리가 원하는 답에 해당하는 H 를 제대로 구하려면, H 의 초기값이 적절하게 주어져야 한다. LM 최소화 방법을 비롯한 대부분의 수식적인 최소화 방법은 글로벌 최소값(Global Minimum)을 찾는 것이 아니고 국부적인 최소값(Local Minimum)을 찾는 것이기 때문에, 초기값이 적절하게 주어지지 않으면 엉뚱한 답을 찾게 된다.

H 의 초기값을 얻으려면 두 이미지 I_0 와 I_1 사이의 대략적인 대응 관계를 알아야 하는데, 사용자로 하여금 I_0 의 위치를 고정시키고 I_1 을 이동시켜 두 이미지가 적당히 겹쳐지도록 하여 I_1 의 초기 위치를 얻거나, 두 이미지 사이의 대응점을 4개 이상 지정해 줄 수 있도록 UI를 제공하는 방법이 있으나, 그렇게 하는 것은 사용자를 매우 번거롭게 하는 것이다. 본 발명에서 자동화를 주안점 중의 하나로 보았기 때문에 그러한 방법을 피하였다.

I_0 의 위치를 고정시키고 I_1 을 이동하여 두 이미지가 잘 겹쳐지는 I_1 의 위치를 찾으려면 잘 겹쳐진 정도를 측정하는 방법이 먼저 정의되어야 한다. 두 이미지가 얼마나 잘 겹쳐졌는지 측정하는 방법으로는 여러 가지가 있는데, 수학식17과 같이 현

재의 위치에서 두 이미지 사이의 대응되는 픽셀의 밝기 차의 자승의 합으로 표현하는 것을 **SSD**(Sum of Squared Difference)라고 한다. SSD 값이 작을수록 두 이미지는 더 잘 겹쳐졌다고 볼 수 있으나, 이 방식은 두 이미지의 전체적인 밝기 및 대조(Contrast)의 차이가 큰 경우에는 제대로 동작하지 않을 수도 있다는 문제점을 가지고 있다.

전체적인 밝기 및 대조의 차이에 상관없는 방법으로, 두 이미지 사이의 **NCC**(Normalized Cross-Correlation)값을 측정하는 방법이 있는데, NCC는 하기의 수학식29 및 수학식30과 같이 정의된다.

수학식 29

$$NCC(x_t, y_t) =$$

$$\frac{\frac{1}{N_R} \sum_{(x,y) \in R} (I_0(x,y) - \bar{I}_0)(I_1(x-x_t, y-y_t) - \bar{I}_1)}{\sqrt{\frac{1}{N_R} \sum_{(x,y) \in R} (I_0(x,y) - \bar{I}_0)^2} \sqrt{\frac{1}{N_R} \sum_{(x,y) \in R} (I_1(x-x_t, y-y_t) - \bar{I}_1)^2}}$$

수학식 30

$$\bar{I}_0 = \frac{1}{N_R} \sum_{(x,y) \in R} I_0(x,y), \quad \bar{I}_1 = \frac{1}{N_R} \sum_{(x,y) \in R} I_1(x-x_t, y-y_t).$$

여기에서 R은 두 이미지 I₀와 I₁이 겹치는 영역을 나타내고, N_R은 R에 속하는 픽셀 수를 나타내며, (x_t,y_t)는 I₁의 이동량을 나타낸다. NCC의 값이 클수록 I₀와 I₁은 잘 겹쳐졌다고 볼 수 있다.

모든 픽셀의 위치에 대하여 NCC 값을 측정하여 NCC 값이 가장 큰 곳에 해당하는 (x_t,y_t)를 찾는 것을 템플리트 매칭이라고 한다. 현실적으로 사진 촬영 시, 카메라의 회전 방향이 항상 일정하고 수평 회전이라는 점을 감안하면, 가장 큰 NCC 값을 가지는 위치를 찾는 데 있어서 탐색 영역을 상당량 줄일 수 있다. 따라서 템플리트 매칭의 수행 시, 이러한 점들이 고려되어야 한다.

템플리트 매칭에 의하여 찾은 위치를 (x₀,y₀)라고 하면, H의 초기값 H₀는 하기의 수학식31과 같이 주어질 수 있다.

수학식 31

$$H_0 = \begin{bmatrix} 1 & 0 & x_0 \\ 0 & 1 & y_0 \\ 0 & 0 & 1 \end{bmatrix}.$$

위에서 제시한 방법을 원래 크기의 이미지에 직접 적용하면 다음과 같은 문제들이 발생한다.

즉, 초기값 H₀와 정답에 해당하는 H 사이의 격차가 여전히 크기 때문에 원하는 답으로 수렴하지 않으며, 템플리트 매칭이나 LM 최소화의 수행 시, 계산량이 너무 많아진다.

이러한 문제점들을 해결하기 위하여 피라미드(Image Pyramid)를 이용한다. 이미지 피라미드는 주어진 원 이미지 및 그 이미지를 축소해서 만든 이미지들로 구성된다. 원 이미지 I의 이미지 피라미드를 구성하는 이미지들의 집합을 Pyramid(I) = {I⁽⁰⁾, I⁽¹⁾, ..., I⁽ⁿ⁾}이라고 한다면, I는 하기의 수학식32와 같이 표현된다.

수학식 32

$$I^{(i-1)} = \text{Reduce}(I^{(i)}), \quad I^{(n)} = I$$

여기에서 Reduce()는 이미지를 축소하는 연산자(Operator)인데, 보통 0.5배로 줄인다.

효율적인 템플리트 매칭 및 LM 최소화를 수행하기 위하여, I₀와 I₁의 이미지 피라미드 Pyramid(I₀)와 Pyramid(I₁)을 생성한다.

이어서, $I_0^{(0)}$ 와 $I_1^{(0)}$ 부터 시작하여 차례대로 템플릿 매칭 또는 LM 최소화를 수행한다. 이때 $I_0^{(i)}$ 와 $I_1^{(i)}$ 의 템플릿 매칭 또는 LM 최소화 수행 시, $I_0^{(i-1)}$ 와 $I_1^{(i-1)}$ 로부터 나온 결과를 초기값으로 활용한다.

이러한 작업을 반복하면, 피라미드의 밑바닥에 가까워짐에 따라 보다 정확한 답을 찾게 된다. 템플릿 매칭의 수행 시, 탐색 영역을 이전 단계에서 구한 위치의 근방(몇 픽셀 이내)으로 설정함으로써 계산량을 상당히 줄일 수 있다.

또한, LM 최소화를 수행할 경우를 고려해 보자면, $I_0^{(0)}$ 와 $I_1^{(0)}$ 에 적용할 경우에는 이미지의 크기가 상당히 작은 관계로 수학식31의 H_0 를 H 의 초기값으로 사용해도 잘 수렴이 됨을 관찰할 수 있다.

또한, 피라미드의 이전 단계에서 나온 결과를 다음 단계의 초기값으로 활용함으로써, 매우 빨리 수렴이 됨을 관찰할 수 있다. 따라서 이미지 피라미드를 활용하면 LM 최소화를 수행할 때 원하는 답으로 수렴이 잘 될 뿐만 아니라 계산량이 상당히 줄어드는 효과가 있다.

또한, 본 발명에서는 카메라 틸트(Camera tilt)를 고려한 이미지 레지스트레이션 방법을 제공하는 바, 이하에서는 카메라 틸트를 고려한 이미지 레지스트레이션 방법에 대해 살펴본다.

촬영 시, 카메라의 틸트 각(지면과 카메라가 바라보는 방향이 이루는 각)이 존재하면 이웃하는 사진들 사이에 회전 각(Rotation angle)이 존재하게 된다.

실험에 의하면, 회전 각이 비교적 큰 경우에는 수학식31의 H_0 를 초기값으로 사용하였을 때, 이미지 피라미드를 채용한다 하더라도 제대로 수렴하지 않는 경우가 발생하였다. 이러한 문제를 해결하려면 이미지 사이에 존재하는 회전 각을 H_0 에 반영해야 한다. I_0 와 I_1 사이의 회전 각 Θ_0 가 주어지면, H_0 는 하기의 수학식33과 같이 수정될 수 있다.

$$\text{수학식 33}$$

$$H_0 = \begin{bmatrix} \cos\theta_0 & -\sin\theta_0 & x_0 \\ \sin\theta_0 & \cos\theta_0 & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

회전 각 Θ_0 를 얻기 위하여, 하기의 수학식34 및 수학식35에 도시한 바와 같이 비용 함수 e 를 LM 최소화 방법을 이용하여 최소화한다.

$$\text{수학식 34}$$

$$e(\theta, x_i, y_i) = \frac{1}{N_R} \sum_{(x,y) \in R} (I_1(x, y) - I_0(x, y))^2,$$

$$\text{수학식 35}$$

$$\hat{x} = x \cos\theta - y \sin\theta + x_i, \quad \hat{y} = x \sin\theta + y \cos\theta + y_i,$$

이 때, 파라메타들의 초기치는 (Θ, x_t, y_t) 로 설정하고, 수렴이 원활하게 되도록 하기 위하여 이미지 피라미드를 도입한다.

또한, 본 발명에서는 페어 와이어 국부적인 이미지 레지스트레이션(Pairwise local image registration) 방법을 제공한다.

카메라 내부 파라메타인 초점거리 f 와 방사 왜곡 계수 K 가 주어지면, 그것들을 활용하여 입력으로 주어진 사진렬 $\{I_1, I_2, \dots, I_n\}$ 의 각 이미지 쌍 $\{I_i, I_{i+1}\}$ 에 대하여 이미지 레지스트레이션을 수행하여 그들 사이의 호모그래피 $H_{i(i+1)}$ 을 구한다.

이미지 레지스트레이션 수행 시, 카메라 틸트를 고려한 이미지 레지스트레이션 방법2을 그대로 적용할 수도 있으나, 그럴 경우에 각각의 이미지 쌍에 대하여 서로 다른 f 및 K 값을 갖는 문제가 발생한다.

따라서, 기존에 구한 f와 K는 공통적으로 사용하고, 나머지 파라메타의 값들을 구하는 방식을 채용해야 한다.

우선, 앞 항에서 구한 K 값을 이용하여 렌즈 왜곡이 없는 이미지를 얻는다. 그런 후, 4-파라메타 호모그래피에 기반을 둔 이미지 레지스트레이션을 수행한다. 4-파라메타 호모그래피에 포함되어 있는 4개의 파라메타는 초점거리 f와 카메라의 회전각 Ψ , Θ , Φ 로 구성되어 있는데, f는 주어진 값으로 고정시키고 카메라의 회전각만을 구하게 된다. 즉, 수학식36 및 수학식37에 도시된 비용 함수 e를 최소화하는 Ψ , Θ , Φ 를 구한다.

수학식 36

$$e(\psi, \theta, \phi) = \frac{1}{N_R} \sum_{(x, y) \in R} (I_{i+1}(\hat{x}, \hat{y}) - I_i(x, y))^2,$$

수학식 37

$$\begin{bmatrix} \hat{x} \\ \hat{y} \end{bmatrix} = \begin{bmatrix} \hat{u}/\hat{w} \\ \hat{v}/\hat{w} \end{bmatrix}, \begin{bmatrix} \hat{u} \\ \hat{v} \\ \hat{w} \end{bmatrix} = KR_{i(i+1)}K^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, K = \begin{bmatrix} f & 0 & x_0 \\ 0 & f & y_0 \\ 0 & 0 & 1 \end{bmatrix}.$$

여기에서 회전 행렬 $R_{i(i+1)}$ 는 수학식4와 같이 주어진다.

위에서 주어진 비용 함수를 최소화하려면, Ψ , Θ , Φ 의 초기값들이 적절히 주어져야 한다. 우선 Ψ 와 Θ 의 초기값을 얻기 위하여 주어진 이미지 쌍 $\{I_i, I_{i+1}\}$ 을 원통 맵핑하여 얻은 이미지 $\{\Gamma_i, \Gamma_{i+1}\}$ 을 가지고 템플릿 매칭을 수행한다. 그런 후, Φ 의 초기값을 얻기 위하여 Γ_i 와 Γ_{i+1} 사이의 회전각을 얻는다.

평면 투영에 의하여 얻어진 이미지는 카메라 파라메타가 주어지면 원통 투영 이미지 또는 구 투영 이미지로 변환이 가능하다. 평면 투영에 의하여 얻어진 이미지 I에 대응하는 카메라의 교정 행렬과 회전 행렬을 각각 K와 R이라고 하면, I에 속하는 임의의 픽셀 $\mathbf{x}=(x,y)$ 을 통과하는 광선(Optical Ray)의 방향 벡터 $\mathbf{X}=(X,Y,Z)$ 는 하기의 수학식38과 같이 주어진다.

수학식 38

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = R^{-1}K^{-1} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

반지름이 1인 원통의 중심이 카메라의 투영 중심과 일치하도록 위치해 있다고 가정하자. 그러면, 광선 \mathbf{X} 와 원통이 만나는 교점 $\mathbf{v}=(\theta, v)$ 을 생각할 수 있는데, 이 점이 원통 투영 이미지 I에서 픽셀 $\mathbf{x} \in I$ 에 대응하는 점이다. 이 때 θ 는 Y축과 광선을 포함하는 평면이 YZ 평면과 이루는 각이고, \mathbf{v} 는 교점의 Y축 좌표로서, 하기의 수학식39와 같이 주어진다.

수학식 39

$$\theta = \tan^{-1}\left(\frac{X}{Z}\right), v = \frac{Y}{\sqrt{X^2 + Z^2}}.$$

반지름이 1인 구에 투영하는 경우에, 광선 X가 구와 만나는 교점의 좌표 (θ, ϕ) 는 하기의 수학식40과 같이 주어진다.

수학식 40

$$\theta = \tan^{-1}\left(\frac{X}{Z}\right), \phi = \tan^{-1}\left(\frac{Y}{\sqrt{X^2 + Z^2}}\right).$$

여기에서 ϕ 는 광선이 XY 평면과 이루는 각이다.

카메라가 Y축을 중심으로 완벽하게 수평 회전을 한 경우(즉, 카메라의 회전각 Ψ 와 Θ 가 0인 경우), 카메라의 회전각을 $\Delta\Theta$ 라고 하고, 회전하기 전에 얻는 원통 투영 이미지를 Γ_0 , 회전한 후에 얻는 이미지를 Γ_1 이라고 하면, 수학식41과 같은 관계가 성립함을 알 수 있다.

수학식 41

$$\hat{I}_0(\theta, v) = \hat{I}_1(\theta - \Delta\theta, v)$$

즉, I'_1 은 I'_0 을 단순히 쉬프팅하여 얻을 수 있다. 따라서, 카메라가 Y축을 중심으로 수평 회전만을 하여 사진을 촬영했을 경우에는 카메라의 초점거리 f 만 알면 각 사진의 원통 투영 이미지를 만든 후, 템플릿 매칭을 통하여 이웃하는 이미지 사이의 쉬프팅 양을 구함으로써 이미지 레지스트레이션이 가능하게 된다. 각 사진의 원통 투영 이미지를 구할 때, 수학식38에서 $R = I$ 로 놓는다

본 발명에서는 좀 더 일반적인 카메라의 회전을 고려하였기 때문에, 위와 같은 방식으로 이미지 레지스트레이션을 수행하지 않고, 위에서 언급한대로 템플릿 매칭을 하여 얻은 결과를 수학식36의 비용 함수를 최소화하기 위한 파라메타의 초기치로 활용한다.

이웃하는 두 이미지 $\{I_i, I_{i+1}\}$ 의 원통 투영 이미지 $\{I'_i, I'_{i+1}\}$ 를 템플릿 매칭하여 얻은 I'_{i+1} 의 쉬프팅 양을 $(\Delta\theta, \Delta v)$ 라고 하면, 카메라의 X축 및 Y축으로의 회전각을 나타내는 Ψ, θ 의 초기치, Ψ_0 와 θ_0 는 하기의 수학식42와 같이 주어진다.

수학식 42

$$\psi_0 = \tan^{-1} \Delta v, \quad \theta_0 = \Delta \theta.$$

카메라의 Z축으로의 회전각 ϕ 의 초기값으로는 I'_i 와 I'_{i+1} 사이의 회전각을 사용하는데, 구하는 방식은 카메라 틸트를 고려한 이미지 레지스트레이션 방법에서 수학식34를 최소화하여 구하는 방식과 동일하다.

카메라 회전각의 초기치가 주어지면, 수학식36의 비용 함수 e 를 LM 최소화 방법을 이용해 최소화함으로써 정확한 카메라의 회전각을 구한다. 위에서 제시한 방법은 이미지 피라미드를 도입하여 수렴성 및 속도를 향상시킨다.

제시한 방법으로 이미지 레지스트레이션을 수행하였을 경우에, 각 이미지 쌍에 대하여 8-파라메타 호모그래피를 기반으로 한 경우보다 안정적이고 균일한 결과를 보인다.

아울러, 본 발명에서는 입력 사진들 사이에 존재하는 밝기 차이를 자동 보정하는 방법을 제공한다.

비록 동일한 카메라를 가지고 사진들을 얻었다고 하더라도 사진들 사이에 전체적인 밝기(Brightness) 및 대조(Contrast)의 차이가 존재하기 마련이다. 이러한 차이를 보정하지 않고 파노라마 이미지를 생성하면, 접합 부분에서 부자연스러운 경계가 발생하게 된다.

어떠한 이미지의 전체적인 밝기 및 대조는 픽셀들의 밝기의 평균값과 표준편차로 표현될 수 있다. 이웃하는 두 이미지 I_i 와 I_{i+1} 사이에 존재하는 밝기 차이를 보정하기 위하여, 서로가 겹치는 영역에서 픽셀들의 밝기의 평균값과 표준편차를 구한다. I_i 와 I_{i+1} 사이의 호모그래피가 주어지면 서로가 겹치는 영역을 찾을 수 있다. 겹치는 영역에서 I_i 의 밝기의 평균값과 표준편차를 m_i^r 과 σ_i^r 이라고 하고, I_{i+1} 의 밝기의 평균값과 표준편차를 m_{i+1}^l 과 σ_{i+1}^l 이라고 한다. 두 이미지의 밝기의 평균값과 표준편차를 같게 만들려면, I_{i+1} 의 픽셀 값을 수학식43 및 수학식44와 같이 보정해야 한다.

수학식 43

$$I'_{i+1} \leftarrow a_{i+1} I_{i+1} + b_{i+1}$$

수학식 44

$$a_{i+1} = \frac{\sigma_i^r}{\sigma_{i+1}^l}, \quad b_{i+1} = m_i^r - a_{i+1} m_{i+1}^l.$$

사진열 $\{I_1, I_2, \dots, I_n\}$ 이 주어지면, 이웃하는 이미지 쌍 $\{I_i, I_{i+1}\}$ 에 대하여 I_{i+1} 의 픽셀 밝기를 수학식43을 이용하여 보정하는 것을 순차적으로 행하는 경우에 대하여 고려해 본다.

360° 파노라마 이미지를 생성하는 경우가 아니라면 특별한 문제가 없으나, 360° 파노라마 이미지를 생성하는 경우에는 처음 이미지와 마지막 이미지를 연결해야 하므로 I_n 과 I_1 의 밝기 차이도 보정이 되어야 하는데, 단순히 순차적으로 보정을 하는 경우에는 I_n 과 I_1 의 밝기 차이가 제대로 보정이 된다고 보장할 수 없다.

360° 파노라마 이미지를 생성하는 경우, 밝기 보정이 순환하는 체인(Chain)의 형식을 띄기 때문에 한번의 순차적인 보정으로는 문제를 해결할 수 없다. 본 발명에서는 하기의 수학식45 내지 수학식47에서 나타낸 바와 같이 비용 함수를 최소화 하는 a_1 과 b_1 을 구함으로써 문제를 해결한다.

수학식 45

$$e(a_1, b_1) = (\sigma_n^r - \sigma_1^l)^2 + (m_n^r - m_1^l)^2$$

수학식45에 나타난 e의 값을 구하는 방법은 다음과 같다. 우선 현재 주어진 $m_n^r, \sigma_n^r, m_1^l, \sigma_1^l$ 값들을 이용하여 수학식 46과 같이 a_1 과 b_1 을 구한다.

수학식 46

$$a_1 = \frac{\sigma_n^r}{\sigma_1^l}, \quad b_1 = m_n^r - a_1 m_1^l$$

구한 a_1 과 b_1 을 이용하여 $m_1^l, \sigma_1^l, m_1^r, \sigma_1^r$ 을 수학식47과 같이 업데이트한다.

수학식 47

$$m_1^l \leftarrow a_1 m_1^l + b_1, \quad \sigma_1^l \leftarrow a_1 \sigma_1^l, \quad m_1^r \leftarrow a_1 m_1^r + b_1, \quad \sigma_1^r \leftarrow a_1 \sigma_1^r$$

업데이트된 m_1^r, σ_1^r 과 기존의 m_2^l, σ_2^l 을 이용하여 a_2 와 b_2 를 구한다. 이것을 반복하여 m_n^r 과 σ_n^r 을 구한 후, e의 값을 구한다.

비용 함수 e를 최소화하는 a_1 과 b_1 을 구하면, 나머지 a_i 와 b_i 들을 위에서 언급한 방식으로 구하고, 수학식43을 이용하여 이미지의 밝기를 보정한다.

이로써, 본 발명은 360° 파노라마 이미지를 생성할 수 있다.

360°로 사진을 촬영하였을 경우, 카메라 파라메타 값들을 완벽하게 구했다면, 앞서 실시한 원통 투영에 의하여 I_n 과 I_1 은 자동적으로 오차 없이 잘 겹쳐질 것이다. 이론적으로 이야기하자면, I_n 과 I_1 사이의 호모그래피로부터 구한 카메라의 회전 행렬을 R_{n-1} 이라고 하였을 때, R은 하기의 수학식48과 같이 표현된다.

수학식 48

$$R = R_{n1} R_{(n-1)n} \dots R_{12}$$

즉, 수학식48에 도식된 R은 회전 행렬들이 오차 없이 구해졌을 경우 완벽하게 일치 함수(Identity Matrix)가 되어야만 한다. 그러나, 어떠한 경우에서든 계산상의 오차는 포함되기 마련이기 때문에, 단순히 페어 웨어 국부적인 이미지 레지스트레이션 만으로는 360° 파노라마 이미지를 얻을 수 없다. 이러한 문제를 해결하기 위하여 수학식49에서와 같이 비용 함수 e를 최소화하는 글로벌 이미지 레지스트레이션 문제를 생각해 볼 수 있다.

수학식 49

$$e(K, R_{12}, \dots, R_{(n-1)n}, R_{n1}) = \frac{1}{N_n} \sum_{(x,y) \in R_n} (I_1(x,y) - I_n(x,y))^2$$

$$+ \sum_{i=1}^{n-1} \frac{1}{N_{R_i(x,y)}} \sum_{(x,y) \in R_i} (I_{i+1}(x,y) - I_i(x,y))^2.$$

그러나, 이 방법은 계산량이 너무 많고, 정확한 360° 파노라마 이미지 생성을 보장하지는 못한다.

본 발명에서는 보다 간단하고 빠른 방법으로 360° 파노라마 이미지를 생성한다. 상기한 과정을 통해 구한 파노라마 이미지를 I_p 라고 하면, 360° 파노라마 이미지를 생성하는 방법은 다음과 같다.

제1단계: I_p 에서 I_1 이 맵핑된 영역을 R_1 이라고 하면, I_p 로부터 R_1 에 속하는 픽셀들만을 추출하여 이미지 I_1 을 만든다.

제2단계: I_p 에서 I_n 이 맵핑된 영역을 R_n 이라고 하면, R_n 내에서 I_p 와 I_1 의 템플릿 매칭을 수행한다. I_1 의 (0,0)에 대응하는 I_p 의 위치를 (x_m, y_m) 이라고 한다.

제3단계: I_1 의 (0,0)의 지점이 I_p 상에서 (x_m, y_m) 에 오도록 위치시킨 후, 이미지 블렌딩(Image Blending)을 수행하여 이미지 I_p' 을 만든다.

제4단계: I_1 의 중심점의 좌표를 (x_c, y_c) 라고 하면, I_p' 으로부터 두 수직선 $x = x_c$ 와 $x = x_m + x_c$ 사이의 영역을 취하여 이미지 I_p'' 을 만든다.

제5단계: I_p'' 을 수직 방향으로 절단(Shearing, Cropping)하여 I_p'' 의 양 끝이 정확하게 일치하도록 한다.

엄청난 양의 디지털 카메라가 현재 보급 되고 있고, 이제 필름 카메라는 그 자리를 잃어가고 있는 실정이다. 또한, 인터넷의 엄청난 보급으로 인하여 디지털 카메라와 인터넷 혹은 콘텐츠는 상호 보완적으로 발전하고 있다. 전술한 바와 같이 이루어지는 본 발명으로 인하여 로테이터와 같은 특수 장비 없이 삼각대만을 이용해서 고품질의 360도 파노라마 이미지를 생성할 수 있고, 렌즈의 초점 거리에 대한 정보를 필요로 하지 않으며, 광각 렌즈를 사용할 때 나타나는 이미지 왜곡 현상을 자동으로 보정할 수 있으며, 카메라가 지면과 수평을 이루지 않은 상태로 촬영된 사진들도 붙이는 것이 가능하며, 입력 사진들 사이에 존재하는 밝기의 차이를 자동으로 보정하며, 원통(Cylinder) 및 구(Sphere) 투영 파노라마 이미지를 생성할 수 있음을 실시예를 통해 알아 보았다.

본 발명의 기술 사상은 상기 바람직한 실시예에 따라 구체적으로 기술되었으나, 상기한 실시예는 그 설명을 위한 것이며 그 제한을 위한 것이 아님을 주의하여야 한다. 또한, 본 발명의 기술 분야의 통상의 전문가라면 본 발명의 기술 사상의 범위 내에서 다양한 실시예가 가능함을 이해할 수 있을 것이다.

발명의 효과

상술한 바와 같은 본 발명은, 디지털 카메라를 가지고 이미지 기반의 가상공간을 특별한 기술 없이도 쉽고 빠르게 만들 수 있도록 하여 보다 많은 수의 이미지 기반 가상 공간 콘텐츠를 만들고, 이를 산업 전반에 확대시킬 수 있는 효과가 있다.

(57) 청구의 범위

청구항 1.

360° 방향으로 복수의 사진을 촬영하는 단계;

촬영된 상기 복수의 사진으로부터 선택된 임의의 이웃하는 사진 한 쌍을 이용하여 카메라의 초점거리와 왜곡률 및 두 이미지 사이의 회전각으로 이루어진 파라메타를 산출하는 단계;

상기 산출된 파라메타를 이용하여 각각의 이웃하는 사진 쌍으로부터 이미지 사이의 상대적인 카메라 회전각들을 산출하는 단계;

상기 이미지 사이의 상대적인 회전각들을 이용하여 각 이미지에 대응하는 카메라 파라메타를 계산하는 단계;
 상기 카메라 파라메타를 이용하여 이웃하는 이미지 사이의 밝기 차이를 보정하는 단계;
 밝기 차이가 보정된 각 이미지를 단일 원통에 투영하는 단계;
 상기 투영 결과, 겹치는 영역에서 이미지 블렌딩을 실시하여 미완성 파노라마 이미지를 형성하는 단계;
 사진 열의 첫 번째 이미지를 이용하여 원통 투영 이미지를 생성하는 단계; 및
 상기 원통 투영 이미지와 상기 미완성 파노라마 이미지를 이용하여 360° 파노라마 이미지를 생성하는 단계
 를 포함하는 360° 파노라마 이미지 생성 방법.

청구항 2.

제 1 항에 있어서,

상기 파라메타를 산출하는 단계는,

상기 임의의 이웃하는 사진 한 쌍에 대해 템플릿 매칭(Template matching)을 통한 두 이미지 사이의 트랜스레이션(Translation) 양을 계산하는 단계와,

LM(Levenberg-Marquardt) 최소화를 통한 두 이미지 사이의 회전각을 계산하는 단계와,

LM 최소화를 통한 렌즈의 왜곡률 및 두 이미지 사이의 8-파라메타 호모그래피(Homography)를 계산하는 단계와,

카메라 자동 교정(Camera auto-calibration)을 통한 카메라의 초점거리를 계산하는 단계를 포함하는 것을 특징으로 하는 360° 파노라마 이미지 생성 방법.

청구항 3.

제 1 항에 있어서,

상기 이미지 사이의 상대적인 카메라 회전각들을 산출하는 단계는,

상기 산출된 파라메타를 이용하여 각각의 이웃하는 사진 쌍으로부터 렌즈 왜곡이 보정된 이미지를 생성하는 단계와,

상기 렌즈 왜곡이 보정된 이미지를 이용하여 원통 투영 이미지를 생성하는 단계와,

이미지 회전 후, 템플릿 매칭을 통한 두 이미지 사이의 트랜스레이션 양을 계산하는 단계와,

LM 최소화를 통한 두 이미지 사이의 3-파라메타 호모그래피를 계산하는 단계를 포함하는 것을 특징으로 하는 360° 파노라마 이미지 생성 방법.

청구항 4.

제 1 항에 있어서,

상기 360° 파노라마 이미지를 생성하는 단계는,

상기 원통 투영 이미지와 상기 미완성 파노라마 이미지를 이용하여 템플리트 매칭을 통한 두 이미지 사이의 트랜스레이션 양을 계산하는 단계와,

이미지 합체 후 절단하는 단계를 포함하는 것을 특징으로 하는 360°파노라마 이미지 생성 방법.

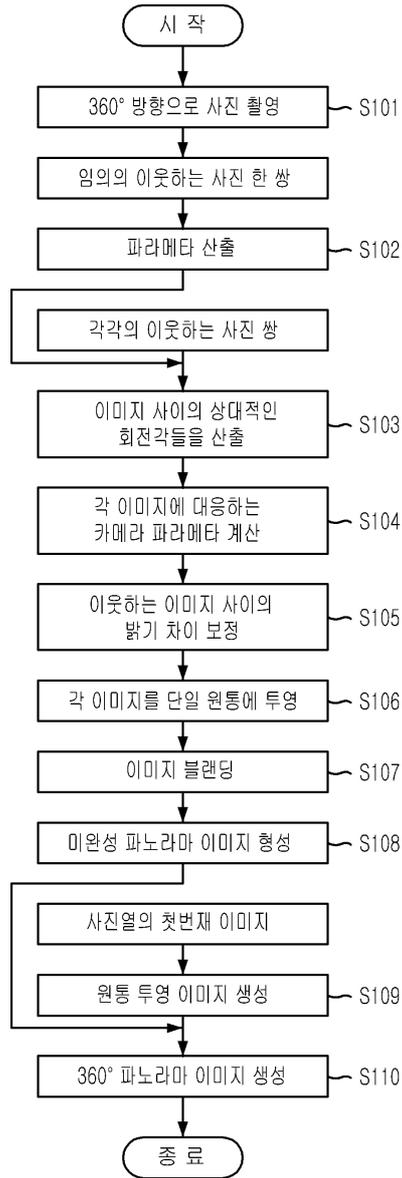
청구항 5.

제 2 항 또는 제 3 항에 있어서,

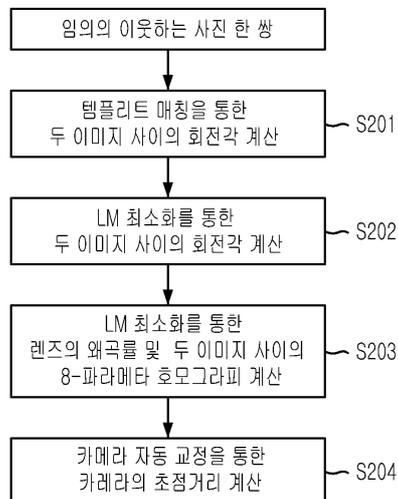
상기 템플리트 매칭 및 상기 LM 최소화의 수행 시 계산량을 줄이기 위해 이미지 피라미드를 이용하는 것을 특징으로 하는 360°파노라마 이미지 생성 방법.

도면

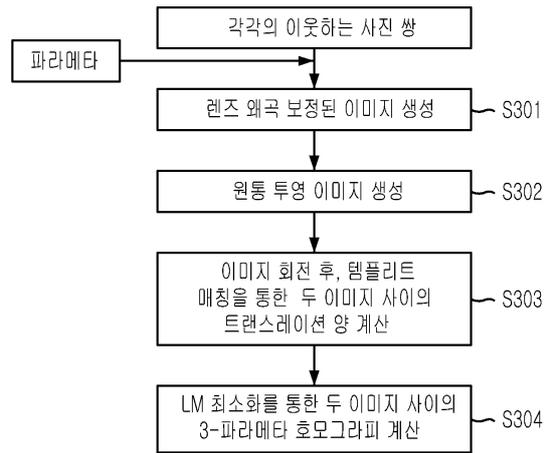
도면1



도면2



도면3



도면4

