

2차원 전기영동 영상의 스팟 정합을 위한 Landmark 스팟쌍의 검출 Detection of Landmark Spots for Spot Matching in 2DGE

한찬명¹, 석수영¹, 윤영우^{2*}
Chan-Myeong Han¹, Soo-Young Suk¹, Young-Woo Yoon^{2*}

<Abstract>

Landmark Spots in 2D gel electrophoresis are used in many methods of 2DEG spot matching. Landmark Spots are obtained manually and it is a bottle neck in the entire protein analysis process. Automated landmark spots detection is a very crucial topic in processing a massive amount of 2DGE data. In this paper, Automated landmark spot detection is proposed using point pattern matching and graph theory. Neighbor spots are defined by a graph theory to use and only a centered spot and its neighbor spots are considered for spot matching. Normalized Hausdorff distance is introduced as a criterion for measuring degree of similarity. In the conclusion, the method proposed in this paper can get about 50% of the total spot pairs and the accuracy rate is almost 100%, which the requirements of landmark spots are fully satisfied.

Keywords : Landmark spot, Neighbor spot, Local spot matching, Delaunay triangulation, Hausdorff distance

1. 2차원 단백질 전기영동

사람을 포함한 40여종의 지놈프로젝트(genome project)이후 유전자 서열정보를 성공적으로 해독하였지만 유전자 서열정보만을 가지고 생명현상을 설명하고 질병과 관련된 유전자를 발굴하는 것에 많은 한계가 있다는 것이 밝혀졌다¹⁾.

유전자는 mRNA와 단백질 발현이라는 복잡한 기전을 가지므로 단백질의 발현과 단백질간의 상호작용을 밝히는 것이 생명현상과 질병을 연구하는 데 중요한 학문으로 부각되었다. 포괄적으로 단백질의 기능을 연구하는 학문을 단백질체학(proteomics)이라고 하는데 단백질체학의 가장 기본이 되는 절차는 세포내에 포함된 단

백질의 종류를 분별하는 것이며 이러한 목적으로 가장 많이 이용되는 것이 2D 단백질 전기영동 방법이다.

1.1 2차원 단백질 전기영동의 원리

전기영동(electrophoresis)은 전기장 안에서 하전된 입자(이온)가 양극 또는 음극 쪽으로 이동하는 현상을 말한다. 이 때 이동하는 속도는 입자의 전하량, 크기와 모양, 용액의 pH와 점성도, 용액에 있는 다른 전해질의 농도와 이온의 세기, 지지체의 종류 등 여러 가지 요인에 의해 결정된다. 그러므로 전기영동법은 아미노산, 뉴클레오티드, 단백질과 같은 하전된 물질들을 분리하거나 분석하는데 매우 효과적인 수단으로

¹정희원, 경북IT융합산업기술원

²교신저자, 영남대학교, 컴퓨터공학과 교수, 工博
E-Mail : ywoon@ynu.ac.kr

¹Gyungbuk Institute of IT Convergence Industry Technology

²Corresponding Author, Department of Computer Engineering, Yeungnam Univ., Prof., Ph.D.

이용된다. 특정한 신체의 세포조직은 수많은 단백질로 이루어져 있다. 신체로부터 채취한 세포조직을 특수한 용액에 담구어 단백질들을 서로 분리하고 이 용액을 점성이 있는 겔 위에 투입하고 양쪽에 전압을 걸어주면 일차적으로 전기장이 형성되어 단백질 분자가 가지는 전하량의 차이에 의하여 단백질 분자들이 겔 용액 상을 이동하게 된다. 결과적으로 전기장에 의해 분리된 단백질은 띠의 모양을 가지게 되는데 단백질들은 또 각기 다른 분자량을 가지고 있으므로 중력에 의하여 2차적으로 분리될 수 있다.

두 단계의 과정을 거치면 단백질들은 전하량과 분자량의 차이에 의하여 2차원 겔 용액상에 자신의 전하량과 분자량에 따라 특정한 위치에 분포하게 되고 이 때 단백질이 발견되는 위치는 이 단백질이 어떤 단백질인지 판별할 수 있는 중요한 정보가 된다²⁾. 전기영동이 끝난 겔상의 단백질들은 눈에 잘 띄지 않으므로 염색의 후처리 과정을 거치고 다른 전기영동 결과와 비교하기 위하여 Fig. 1과 같이 이미지 형태로 저장된다.



Fig. 1. 2DGE image.(The black spot represents protein)

1.2 2차원 단백질 전기영동에서 Spot Detection

Fig. 1과 같이 2차원 전기영동의 결과 이미지에서 검은 스팟이 단백질 부분을 나타내는데 각각의 스팟이 각기 다른 단백질을 나타내며 스팟의 개수가 수십에서 수백에 이르고 스팟의 색깔이 배경과 비슷한 것이 존재하므로 인간이 수작업으로 단백질을 분리하거나 다른 전기영

동 결과와 비교하는 것은 거의 불가능하다.

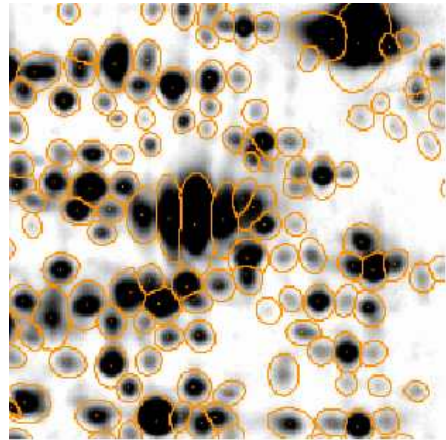


Fig. 2. Detected spot image.

결과 영상에 대하여 영상처리 기법을 적용하여 스팟의 3차원 기하학적 모양으로 배경이 아닌 실제 단백질 위치를 구하게 되는데 이를 스팟의 검출(Spot Detection)이라 한다.

1.3 2차원 단백질 전기영동에서 Spot Matching

질병이 있는 세포조직은 정상 세포조직과 다른 단백질이 발현된다. 따라서 2차원 전기영동 영상에서 어떤 단백질이 출현하였는지를 검색하는 것이 중요한데 스팟이 어떤 단백질인지 판별하기 위해서는 해당 스팟을 추출하여 질량 분석기를 이용하여 분석하여야 한다. 하지만 이 방법은 많은 시간과 높은 비용을 요구하므로 특수한 경우에만 쓰이며 보통은 단백질 분석이 완료된 레퍼런스 이미지(reference gel image)와 새로 획득된 전기영동 이미지를 비교하여 두 영상간의 스팟 매칭(spot matching)을 통한 차이를 분석하는 방법이 쓰인다. 스팟의 매칭은 주로 스팟의 중심 좌표를 이용하여 비교하는데 스팟의 형태는 동일한 세포조직을 여러번 실험하더라도 모두 다르게 획득되므로 주로 스팟의 위치를 이용하는 것이 일반적이다.

스팟의 중심좌표를 이용하면 이것은 계산기하학의 한 부류인 Point pattern matching의 문제가 된다³⁾.

2. 2차원 단백질 전기영동에서 Landmark 스팟쌍의 검출

2.1 Landmark의 검출 필요성

2차원 단백질 전기영동에서 비교 겔 영상 상에 존재하는 스팟을 참조 겔 영상(reference gel image)상에 스팟과 매칭하여 매칭쌍을 생성한다. 2차원 단백질 전기영동의 특성상 동일한 조직의 세포로 동일한 실험을 반복할 때 동일한 단백질 스팟의 위치는 매 실험마다 다르게 도출된다. 이는 2차원 단백질 전기영동의 실험이 수많은 요인에 대하여 영향을 받기 때문이며 가장 큰 원인은 겔 농도의 상태를 일정하게 유지하는 것이 어렵기 때문이다.

더군다나 다른 실험실에서 동일한 세포조직에 대하여 수행된 2차원 단백질 전기영동 영상은 Fig. 3과 같이 큰 차이를 보인다. 2차원 단백질 전기영동 스팟 매칭의 방법 중에는 영상 전역에 분포하는 두드러진 몇 개의 스팟쌍(Landmark)을 사람이 수동으로 지정하여 이들 정보로부터 한 쪽 이미지를 와핑(warping)한 후 두 이미지를 겹쳐 동일한 스팟쌍을 구하는 기법들이 있다.

이 때 Landmark를 지정하는 것은 사람에 의한 수동적인 방법을 사용하기 때문에 대량의 실험을 통한 결과 도출이 불가능하다. 만약 이 부분을 자동화 할 수 있다면 대량의 실험을 모두 자동화 할 수 있으므로 획기적인 방법이 된다.

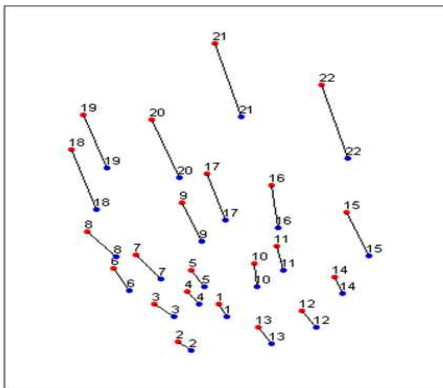


Fig. 3. Local distortion of 2DGE.

2.2 Landmark 스팟쌍의 요건

Landmark의 정보를 이용하여 참조 겔 이미지에 대하여 비교 겔 이미지에서 warping을 수행하며 2차원 단백질 전기영동의 특성상 두 겔 이미지의 왜곡의 특성이 지역마다 차이가 있으므로 Landmark는 이미지 전역에 고루 분포하는 것이 좋다.

두 겔 이미지에서 단백질 변성으로 인하여 단백질 매칭쌍이 존재하는 않는 경우도 있으므로 영상처리에 의한 Landmark 스팟쌍의 정확도 아주 중요하다. 만약 오류에 의하여 스팟쌍이 아닌 스팟들이 쌍으로 매칭되면 전체 스팟쌍 매칭에 치명적인 영향을 미치기 때문이다.

2.3 Landmark 스팟쌍 매칭의 방법

기존의 논문들에서 Landmark 스팟쌍은 사람에 의해 수동적인 방법으로 구해졌다. 사람이 어떠한 방법으로 Landmark 스팟쌍을 찾는지에 대한 고찰은 이 부분에 대한 알고리즘을 고안하는데 큰 도움이 된다.

사람이 Landmark 스팟쌍을 찾는 방법은 사람마다 차이가 있을 수 있겠지만 대부분 가장 스팟의 크기가 크고 두드러진 스팟을 참조 겔 이미지에서 우선적으로 선정하고 그 스팟을 중심으로 주변의 몇 개의 스팟을 확인한 후 비교 겔 이미지에서 대략의 유사한 위치에서부터 시작하여 중심 스팟과 주변의 몇 개의 이웃 스팟의 위치가 유사한 스팟을 검색한다.

이러한 방법으로 인간은 몇 개의 Landmark 스팟을 손쉽게 구해 낼 수 있는데 이것을 알고리즘으로 구현하고자 한다면 다음과 같이 두 가지의 정의가 필요하다. 첫 번째는 이웃스팟의 정의란 무엇인가와 중심 스팟과 이웃스팟의 위치가 유사하다는 것을 어떻게 평가할 수 있겠느냐 하는 것이다.

2.4 이웃스팟의 정의

이웃스팟(Neighbor spot)은 그래프 이론에서 쉽게 정의될 수 있다. 그래프는 정점(vertex)와 두 정점을 잇는 간선(Edge)로 이루어지는데 아래와 같이 정점의 집합 V , 간선의 집합 E 로 표시하며 그래프는 정점과 간선으로 이루어지게 된다.

$$\begin{aligned}
 V &= \{1, 2, 3, 4, 5, 6\} \\
 E &= \{(1,2), \{1,5\}, \{2,3\}, \{2,5\}, \{3,4\}, \{4,5\}, \{4,6\}\} \\
 G &= (V, E)
 \end{aligned}$$

위의 그래프에 대한 수학적 표기는 아래의 Fig. 4와 같이 표현될 수 있다.

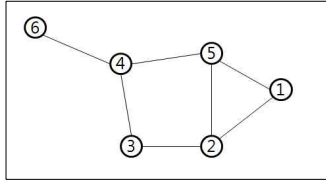


Fig. 4. Graph represent.

위의 그래프에서 정점을 스팟으로 간주할 때 각 스팟에 대한 이웃스팟은 스팟과 스팟간의 간선이 존재하는 것으로 정의할 수 있다. 예를 들면 스팟1의 이웃스팟은 스팟2와 스팟5가 된다. 따라서 각 스팟에 대한 이웃스팟을 아래와 같이 수학적으로 표현할 수 있다.

$$\begin{aligned}
 N_G(v) &= \{u \in G | v u \in E\} \\
 N_G(1) &= \{2,5\} \quad N_G(2) = \{3,5\} \quad N_G(3) = \{2,4\} \\
 N_G(4) &= \{3,5,6\} \quad N_G(5) = \{2,4\} \quad N_G(6) = \{4\}
 \end{aligned}$$

그래프는 그 정의에 따라 많은 그래프가 존재하는데 일단 그래프가 정의되면 자연스럽게 이웃스팟도 그래프의 정의에 의하여 결정된다.

Point pattern matching에서 자주 사용되는 graph는 nearest neighbor graph, Gabriel graph, Delaunay triangulation가 있다.

2.5 이웃 스팟에 의한 지역 매칭(Local matching)

두 켤 이미지에서 각각 중심 스팟과 그 중심 스팟의 이웃스팟들을 가지고 스팟쌍의 일치 여부를 판별하는 것을 이웃 스팟에 의한 지역매칭이라고 한다. 초기의 2차원 단백질 전기영동 스팟 매칭 논문들에서는 모든 스팟들을 전역적으로 매칭하는 방식을 많이 사용하였는데 켤 이미지의 특성상 두 켤 이미지 간의 왜곡이 전

역적(globally)으로 존재함과 동시에 지역적(locally)으로 무수히 많은 왜곡 존재하기 때문에 현재는 지역매칭을 통한 방법이 많이 연구되고 있다. 중심 스팟과 주변 이웃스팟들을 가지고 스팟쌍임을 비교하기 전에 반드시 해야 할 일은 두 켤 이미지에서 비교를 위해 선택된 두 중심 스팟들 간에 지역적 왜곡을 제거하는 일이다.

지역적인 왜곡은 회전(Rotation), 크기 변환(Scaling), 그리고 이동(Transition)으로 Affine 변환을 통해서 제거될 수 있다. 두 개의 스팟쌍이 결정되면 Affine 변환에서의 모든 파라미터를 구할 수 있는데 아래의 그림에서와 같이 중심 스팟쌍은 항상 일치한다고 가정하고 이웃스팟들을 조합으로 선택하여 두 개의 스팟쌍을 결정하고 그것으로부터 Affine 파라미터를 구한다. Affine 파라미터가 결정되면 한 쪽의 중심 스팟과 주변스팟을 Affine 변환한 후 중심 스팟쌍을 기준으로 포개 후 매칭쌍을 결정하고 이를 이용하여 유사도를 평가한다.

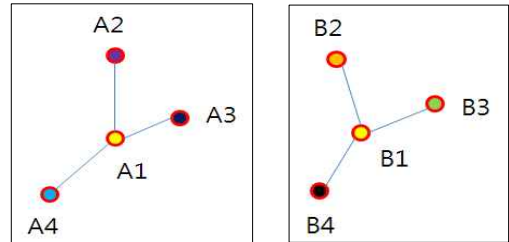


Fig. 5. Separation of the center spot and its neighbor spots using graph.

Fig. 5는 각각 두 켤 이미지의 일부이다. 원으로 표시된 것은 스팟이며 스팟과 스팟 사이에 간선은 이웃의 관계를 나타낸다. A1과 B1은 중심 스팟이며 A2, A3, A4는 A1의 이웃스팟이며 B2, B3, B4는 B1의 이웃 스팟이다. 켤 이미지 A와 켤 이미지 B를 비교하기 위해서는 두 켤 이미지를 A1과 B1을 중심으로 포개보면 되는데 서로 다른 지역적 왜곡으로 인하여 비교가 불가능하므로 A2와 B2를 스팟쌍으로 가정하면 중심 스팟쌍 (A1, B1)과 이웃 스팟쌍 (A2, B2)의 두개의 스팟쌍이 정해진다. 이로부터 Affine parameter를 구한다.

겔 이미지 B를 획득된 Affine parameter를 이용하여 변환 후 두 겔 이미지를 Fig. 6의 (b)와 같이 포개어 비교한다. 2차원 단백질 전기영동에서 극심한 회전이 일어나지 않는다고 가정하면 가능한 두 개의 스팟쌍은 (A1-B1, A2-B2), (A1-B1, A3-B3), (A1-B1, A4-B4)이 된다. 이 세 가지의 경우대하여 각각 Affine parameter가 계산되므로 각각에 대하여 Affine parameter를 적용 후 두 겔 이미지를 포개어 비교한 후 가장 오류가 작은 것은 선정하면 된다.

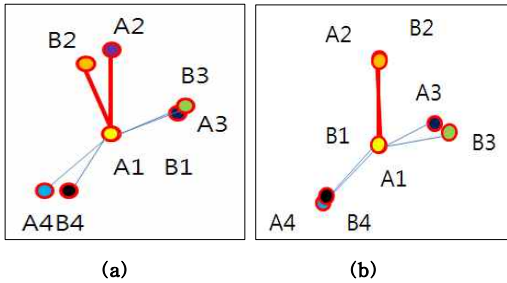


Fig. 6. Removal the local image distortion using affine transformation. (a) Overlapping of the center spot without affine transformation. (b) Overlapping of the center spot with affine transformation using (A4,B4).

2.6 2차원 전기영동에서 스팟간의 유사도 평가

2차원 단백질 전기영동에서 두 겔 이미지에서 서로 매칭되는 스팟쌍을 찾기 위하여 사용할 수 있는 정보는 스팟의 위치 정보밖에 없다. 2차원 단백질 전기영동의 특성상 이런 스팟의 위치정보조차도 지역에 따라 각기 다른 오차가 존재한다. 이러한 특성 때문에 Point pattern matching의 문제 중 2차원 단백질 전기영동의 스팟 매칭문제는 아주 까다로운 문제이다.

두 겔 이미지에는 스팟쌍이 존재하지 않는 스팟들이 존재한다. 이는 특정한 단백질이 질병에 의하여 다른 단백질로 발현되는 차이에서 발생한다. 따라서 정확히 일치하는 스팟쌍이라고 하더라도 이웃스팟이 달라질 수 있다. 유사도는 스팟쌍간의 유클리드 거리를 모두 합한 값에 평균을 구하는 것이 일반적이다. 하지만 앞에서 기술한 바와 같이 스팟쌍이 존재하지 않을 경우도 있기 때문에 Hausdorff distance를 유사도의 척도로 이용하는 것이 타당하다.

3. 2차원 단백질 전기영동에서 Landmark 스팟쌍의 검출 구현

3.1 실험 데이터의 선정

2차원 단백질 전기영동에서 스팟의 검출의 단계를 먼저 수행한 후 검출된 스팟의 중심 좌표를 추출하여 이것을 가지고 스팟 매칭의 단계를 수행한다. 따라서 스팟 매칭에 필요한 정보는 스팟들의 중심좌표만 필요하다.

본 논문에서는 스팟 매칭의 정확성을 객관적으로 평가하기 위하여 본 논문에서 스팟 검출 단계를 생략하고 웹사이트 [http://www.lecb.ncifcrf.gov/2DgelData Sets](http://www.lecb.ncifcrf.gov/2DgelData%20Sets)에서 수록된 Human leukemias 데이터 셋을 사용하였다. 이 셋은 128개의 이미지 쌍에 대하여 스팟매칭을 완료하여 텍스트 파일로 작성해 두었으므로 이를 그대로 이용하여 실험을 수행하였다.

3.2 이웃스팟을 검출을 위한 그래프 생성

중심 스팟 주위에 분포하는 이웃스팟을 정의하기 위하여는 우선 어떤 그래프를 사용할지를 결정하여야 한다. 그래프의 정의에 따라 같은 겔 이미지에서 동일한 스팟에 대하여 서로다른 이웃스팟이 산출된다.

비교하고자 하는 한쌍의 겔 이미지에서 매칭되는 스팟쌍들은 동일한 이웃스팟들을 가지는 것이 바람직하나 2차원 단백질 전기영동의 특성상 예측할 수 없는 지역적 영상 왜곡이 동반되는 것이 일반적이다. 따라서 지역적 영상 왜곡에 대하여 강건한 그래프를 선정하여 매칭되는 스팟쌍에 대하여 가능한 동일한 이웃스팟들을 가지도록 하는 것이 필요하다. 본 논문에서는 지역적 영상왜곡에 대하여 유사한 이웃스팟들을 산출하는 Delaunay triangulation을 이용하였다.

Delaunay triangulation을 점의 집합에서 세 점을 선택하여 세 점에 외접하는 원을 그린 후 그 원 안에 세 점을 제외한 다른 점들이 포함되어 있지 않다면 세 점을 서로 연결하여 삼각형을 구성하여 그래프를 생성한다. 아래의 그림은 정확히 동일한 단백질 스팟들을 포함하는 두 겔 이미지에서 Delaunay triangulation을 이용

하여 생성한 그래프를 나타내었다.

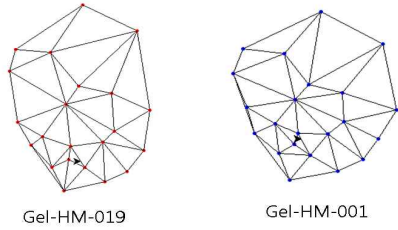


Fig. 7. Resultant graphs of two gel images including the same protein spots using Delaunay triangulation.

3.3 이웃스팟 비교에 의한 스팟 local matching

스팟의 중심 좌표를 이용하여 그래프를 생성하면 이웃스팟의 관계가 결정된다. 일단 참조 겔 영상에서 임의의 한 스팟을 선택하고 비교 겔 영상에서 임의의 한 스팟을 선택한 후 이 두 스팟이 매칭되는 스팟인지 비교하기 위해서 2.4에서처럼 각 두 중심스팟의 이웃스팟만 가지고 비교를 수행하게 된다.

각 겔 이미지에 N개의 스팟이 포함되어 있다면 조합에 의하여 N2번의 비교가 필요한데 비교의 횟수를 줄이기 위하여 먼저 중심스팟의 이웃 스팟의 개수를 비교하는 것이 필요하다. 매칭되는 스팟쌍에서 각 중심스팟의 이웃스팟의 개수는 동일한 것이 일반적이지만 2차원 단백질 전기영동에서는 지역적 영상왜곡에 때문에 이웃 스팟의 개수가 한 개 혹은 두 개 정도 차이가 나므로 이점을 이용하면 비교시간을 줄일 수 있다.

비교하기 전 지역적 영상 왜곡을 해소하기 위하여 2.4에서 설명했듯이 affine 변환을 적용하고 두 중심스팟을 기준으로 두 스팟들의 집합을 정확히 포갠다. 만약 이웃스팟 중에서 매칭되는 스팟이 존재하지 않는다면 그 이웃스팟은 제외하고 매칭쌍이 존재하는 이웃스팟들을 이용하여 Hausdorff 거리를 측정한다.

Hausdorff의 거리가 작으면 유사도가 높다고 간주할 수 있다.

3.4 Hausdorff거리의 정규화

2차원 단백질 전기영동에서는 수많은 지역적 영상 왜곡이 존재한다. 이는 경우에 따라서 다

양한 scale parameter를 산출하게 되는데 scale parameter가 작을수록 동일한 비교라 하더라도 Hausdorff 거리가 작게 산출된다. 그림에서 B는 A를 단순히 확대시킨 경우이다. A와 B의 비교에서 유사도는 두 경우 모두 동일하나 B의 scale factor가 크기 때문에 A가 더 작은 Hausdorff 거리를 가지게 되고 결국 A의 유사도가 더 높은 결과를 산출하게된다. 따라서 이를 고려하지 않고 단순히 Hausdorff 거리만을 가지고 비교하는 것은 타당하지 않다. 따라서 affine 변환에서 scale factor를 고려하기 위하여 Hausdorff 거리를 정규화 하는 것이 필요하다. Hausdorff 거리를 정규화 하는 것은 Hausdorff 거리에 피벗 거리를 나누어 주면 된다. Fig. 8과 같이 피벗 거리는 중심스팟으로부터 affine parameter를 계산할 때 사용된 이웃스팟까지의 거리이다.

3.5 실험결과

3.1에서 언급한 Human leukemias 데이터 셋을 이용하였다. 128개의 이미지 쌍에 대하여 미리 매칭된 스팟쌍과 각 스팟의 중심좌표가 수록되어 있는 텍스트 파일을 이용하였다. 모든 이미지 쌍에는 동일한 수의 스팟들이 포함되어 있으며 평균 22쌍의 스팟이 포함되어 있다. 일단 스팟의 매칭여부를 알고 있기 때문에 실험 후 매칭된 스팟쌍이 올바른지 알 수 있다. 총 2763개의 스팟쌍을 대상으로 실험하였다. 실험 시 정규화된 Hausdorff 거리를 이용 특정값 이상의 Hausdorff 거리에 대해서 잘못된 매칭이 수행된 결과로 간주하였다.

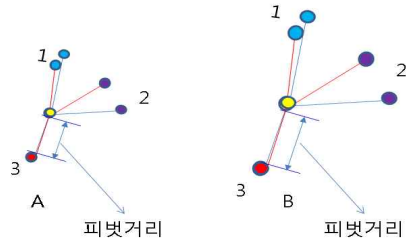


Fig. 8. Normalized Hausdorff distance.

매칭 스팟쌍은 주어진 정규Hausdorff값보다 작은 값을 가지고 매칭된 스팟쌍을 말한다. Hausdorff값이 클수록 매칭 스팟쌍의 늘어나지

만 실제로 스팟쌍이 아닌 false positive의 수가 증가하게 된다.

Table 1. Experimental result

정규Hausdorff값	0.4	0.35	0.3
총스팟쌍	2763	2763	2763
검출 매칭스팟쌍	1430	1409	1377
False 스팟쌍	21	1	0
스팟 검출률	52.51%	51.03%	49.83%
스팟쌍 정확도	98.55%	99.92%	100%

False positive 스팟쌍은 사전에 모든 매칭되는 스팟쌍의 정보를 가지고 있기 때문에 실험에서 도출된 결과와 실제 스팟쌍의 여부를 비교함으로써 판별할 수 있다. 검출률은 총 2763개의 스팟쌍중에서 실험을 통하여 검출된 스팟쌍의 개수를 의미하며 스팟쌍 정확도는 검출된 매칭스팟쌍 중에서 실제로 스팟쌍이 것의 비율을 의미한다.

4. 결론

Landmark 스팟쌍은 두 겹 이미지를 Warping하기 위하여 필요한 스팟쌍을 말하며 전체 스팟쌍중 소수의 스팟쌍으로도 충분하지만 반드시 겹 이미지 전체에 고루 분포하는 것이 좋다. 만약 Landmark 스팟쌍이 false positive의 스팟쌍이라면 이것은 추후 Warping 과정에서 엄청난 부정적인 효과를 가지기 때문에 Landmark 스팟쌍은 반드시 올바르게 매칭된 쌍을 이용하여야 한다. 표에서 스팟 검출률은 52.51%, 51.03% 49.83%로 거의 전체 스팟쌍의 절반의 스팟쌍이 검출된 결과를 보여준다.

이것은 Landmark 스팟쌍으로서 수적인 요건을 충분히 만족시킬 수 있음을 보여주고 검출된 스팟쌍이 겹 이미지 전역에 골루 분포할 확률이 높음을 의미한다. 스팟쌍의 정확도는 각각 98.55%, 99.2%, 100%을 결과를 보이는데 이는 아주 고무적인 결과다. 전술한 바와 같이 false positive 스팟쌍은 이미지 warping에 대하여 상당히 부정적인 영향을 미치므로 반드시 올바르게 매칭된 스팟쌍을 이용하는 것이 아주 중요하다. 실험결과에서 보듯이 스팟쌍의 정확도가 98.5%, 99.2%, 100%라고 하는 것은 이 점을 아

주 잘 충족시켜주고 있다. 실험의 결과로 볼 때 논문에서 제안한 방법은 Landmark 스팟쌍 검출을 자동화하기에 충분한 방법임이라고 할 수 있다.

향후 계획으로서는 더 많은 데이터에 대하여 실험하여 제안된 기법이 실제로 유용함을 증명할 필요가 있다. 실제로 검출된 Landmark 스팟쌍을 이용하여 스팟 매칭을 수행하는 것에 대하여 연구와 본 논문에 제안된 방법에 대하여 스팟 검출율을 더욱 높이는 연구도 병행되어야 할 것이다.

참 고 문 헌

- 1) Wilkins, M.R., K.L. Williams, R.D. Appel, and D.F. Hochstrasser (eds.), *Proteome Research: New Frontiers in Functional Genomics*. Springer, Berlin. Wilkins, M.R., K.L. Williams, R.D. Appel, and D.F. Hochstrasser (eds.), *Proteome Research: New Frontiers in Functional Genomics*. Springer, Berlin.
- 2) Blackstock, W. P. and M. P. Weir. *Proteomics: quantitative physical mapping of cellular proteins*. Trends Biotechnol. (1999)
- 3) J. Park, and P. Kim, "Implementation of embedded software for mobile robot", Proc. of 3rd Asian Embedded Systems Conference, Vol. 2, pp. 100-105, January (2006)
- 4) Link, A.J. *2D Proteome Analysis Protocol in Methods in Biol.*, vol., 112, Humana Press, Totowa, New Jersey (1999)
- 5) E Bettens, P Scheunders, J Sijbers, D Van Dyck, and L Moens. *Automatic segmentation and modelling of two-dimensional electrophoresis gels*. In *International Conference on Image Processing*, volume 2, pp. 665-668, (1996)

(접수:2011.05.24, 수정:2011.06.30, 게재확정:2011.08.24)