

## 음영 분석에 의한 내시경 영상의 3차원 형체 추정에 관한 연구

이태수<sup>o</sup>, 차은종

충북대학교 의과대학 의공학교실

### 3 Dimensional shape estimation by shading analysis of endoscopic image

Tae-Soo Lee<sup>o</sup>, Eun-Jong Cha

Dept. of Biomedical Engineering

College of Medicine, Chungbuk National University

#### 1. 서론

내시경의 역사는 1800년대부터 시작되었지만, 광섬유를 이용한 현대적인 내시경은 1956년 Hirschowitz에 의해 제작된 위십이지장 파이버스코프(fiberscope)로부터 비롯되었으며, 이를 기점으로 해서 본격적인 내시경의 시대가 열리게 되었다. 1983년에는 CCD를 이용한 전자 내시경의 출현으로 또한번의 획기적인 발전이 있었는데, 이로써 컴퓨터를 이용한 영상의 처리, 분석, 및 관리가 용이하게 되었다. 내시경은 인체 내부기관의 정확한 진단 및 치료에 필수적인 기기가 되었지만, 위내시경의 경우 식도를 통해 삽입하는데 따른 환자의 고통이 매우 심하며, 의사가 진단할 때에도 내시경의 끝에 달린 렌즈의 시야 범위가 제한되어 있어, 질 환 부위의 위치를 파악하는 데는 오랜 경험을 필요로 하여, 숙달하는 데는 많은 노력이 드는 힘든 검사이다. 이를 해결하기 위하여 최근에 여러가지 연구가 시도되어 오고 있는데, 그 중 하나는 삼킬 수 있을 정도의 유연성이 있는 삽입튜브의 개발이다. 이 경우 환자의 고통은 어느 정도 해결할 수 있으나, 삽입 튜브 자체의 유연성으로 인해 의사가 진단하고자 하는 부위로 내시경의 끝부분을 정확히 위치시키기가 어려우며, 질환부위가 내시경에 관찰되더라도 어느 부위인지 알기 어렵다. 본 연구에서는 이 문제를 해결하기 위해 내시경 영상의 처리를 통해 위벽면의 대체적인 3차원 형체를 추정할 수 있는 기법을 제안하고자 한다.

#### 2. 내시경 영상의 특성

내시경 영상은 일반적인 영상과 구별되는 고유의 특성을 가지고 있는데, 이들은 내시경 자체의 기기 특성에 의한 것과 대상이 되는 인체 기관의 특성에 의한 것, 두가지로 분류할 수 있다. 먼저 내시경 기기 자체에 의한 영상의 특징으로는 첫째로 광원의 위치와 카메라 렌즈의 위치가 근접하고, 그 방향이 일치한다는 것이다. 따라서 관찰 각도가 90°정도 되는 상용의 전자내시경에서 촬영되는 영상은 그 밝기가 화소가 위치한 벽면의 정면 기술기에 의해 좌우된다고 할 수 있다. 둘째로는 초점심도(depth of

field)가 5-100mm 정도로 제한되어 있어서, 심도를 벗어나는 위치에 있는 벽면의 경우 초점이 제대로 맞지 않아서 영상상의 주름이나 반점 등의 에지가 선명하게 나타나지 않고 불분명하게 보인다는 것이다. 셋째로는 카메라의 끝이 인체 내부로 삽입되므로 그 위치 및 방향에 대한 정량적인 기준을 정할 수 없고, 카메라의 이동에 의한 위치 정보의 추정이 불가능하다는 점이다.

다음은 인체의 기관에 의한 영상의 특성으로서 위벽면을 관찰하는 경우의 특성을 살펴 보면, 첫째로 위의 해부학적인 구조가 오목한(concave) 모양의 대만부와 볼록한(convex) 모양의 소만부 및 아래 부분의 전정부 및 윗부분의 위체부로 구성되어 있는데, 제한된 관찰 각도로 인해 이중의 일부뿐만이 영상으로 구성된다. 둘째로 내시경 검사중의 위는 내부가 비어 있음은 물론 공기를 주입하여 부풀은 튜브 모양이 된다는 것이다. 따라서 그 모양이 급격한 변화를 일으키는 일반적인 영상과 비교할 때 매우 간단한 모양이며, 위벽면에 병변이 있는 경우를 제외하고는 대개 핑크빛을 나타내며, 색조의 변화나 결손이 없는 상태이다 [1]. 셋째로 위는 진단하고 있는 중에도 연동운동에 의한 벽면의 움직임이 계속되고 있으며, 이는 공기가 주입되어 위 벽면이 팽창되어도 정도의 차이일 뿐 무시할 수 없을 정도이다.

살펴본 바와 같은 영상의 특성으로 인해 일반적으로 사용되는 3차원 정보의 추출기법 중에서 내시경에 사용될 수 있는 것은 제한되어 있다.

#### 3. 3차원 재구성 기법

일반적으로 사용되는 3차원 재구성 기법에는 스테레오 차이법 [2], 초점 맞추기법[3], 텍스처 분석법, 음영 분석법[4] 등이 있는데, 본 절에서는 이들 기법에 대한 간단한 소개와 내시경 영상에의 적용 가능성 여부에 대해 논하고자 한다.

##### (1) 스테레오 차이법

이 기법은 가장 일반적으로 사용되는 것으로, 두개의 카메라에 의해 영상을 입력할 때, 그 차이를 검출해서 영상상의 특징점이

위치하는 3차원적 위치를 계산한다. 여기에는 두 카메라의 위치가 렌즈의 광학축과 수직으로 병렬로 있는 경우와 광학축과 같은 방향으로 직렬로 있는 두 경우가 있다. 현재 상용의 내시경은 카메라가 하나이므로 카메라를 광학축 방향으로 이동하여 두 카메라에서 영상을 획득하면 직렬의 배열로 간주할 수 있다.

그러나 두 경우에 모두 카메라의 위치와 방향이 정확히 알려져 있어야 하는데, 내시경에 적용 가능한 직렬 배열의 경우 카메라의 이동거리 및 방향을 정확히 외부에서 제어하기는 매우 어렵다. 따라서 이를 정확히 추정할 수 있는 다른 기법과의 병행없이 직접 적용하기에는 무리가 있을 것으로 판단된다.

(2) 초점 맞추기법

렌즈가 하나인 카메라의 경우에 렌즈의 초점거리와 촬상소자의 위치인 알면, 초점이 정확히 맞는 위치는 렌즈의 공식에 의해 간단히 계산되며, 이보다 가깝거나 먼 경우 초점이 맞지 않아서 예지가 불분명하게 되는데, 이 정도를 계산해서 3차원 정보를 계산할 수 있다. 그러나 내시경의 경우 심도가 5-100mm 정도로 상당히 큰편이며, 예지가 있는 있는 위치에서만 3차원 정보를 계산 가능한 단점으로 인해 이 또한 보조적인 수단 없이 단독으로 위벽면을 재구성하기에는 문제가 있다.

(3) 텍스처 분석법

물체의 표면에 일정한 형태의 텍스처가 있는 경우에 적용할 수 있는 기법으로서, 위벽면의 경우에는 텍스처가 없으므로, 이 기법을 적용하기 위해서는 조명에 격자 무늬가 나타나게 하는 등의 전용 내시경을 제작하여야 하므로 상용 내시경 영상에서는 곤란하다.

(4) 음영 분석법

내시경 사진으로 부터 진단의사가 3차원적인 위의 형태를 머리속으로 추정할 수 있는 것은 위모양에 대한 해부학적인 지식 이외에 영상상에 있는 음영에 의해서도 가능하다. 조명을 받는 부분의 경사가 클수록 음영이 커지며, 경사가 작을 수록 음영이 작아지는 원리를 이용하여 3차원 정보를 추출할 수 있다.

상술한 바와 같이 단독으로 내시경에 적용할 수 있는 기법은 음영분석법 이외에는 많은 문제점을 가지고 있는 것으로 판단되며, 본 논문에서는 내시경 영상의 음영 분석을 통해 3차원적인 위의 벽면을 재구성하고자 한다.

4. 음영 분석을 통한 형태의 재구성

음영분석을 통한 3차원 형태의 재구성 기법은 1970년대에 B.K.P. Horn에 의해 처음 시도된 것으로[5], 이를 이용해서 그날의 형태를 광학적 분석을 통해 추정하였는데, 이후에 Ikeuchi 등에 의해 더욱 발전되었다[6].

Orthographic image projection과 Lambertian illumination을 가정할 때, 내시경 영상 상에서 좌표(x,y) 위치에서의 밝기 E(x,y)는 위벽면에 수직인 벡터에 의해 표시할 수 있다[7]. 3차원 상에서의 z축이 조명방향 및 카메라 렌즈의 광학 방향과 일치하는 내시경 영상에서는 이 수직벡터는 (x,y) 위치에서의 벽면의 z축 좌표, 즉 z(x,y)의 x방향 및 y방향 편미분값으로 나타낼 수 있으며, 이를 함수로 표시하면 식(1)과 같다.

$$E(x,y) = R(z_x(x,y), z_y(x,y)) \tag{식(1)}$$

함수 R은 위벽면의 기울기가 0일때, 즉 렌즈의 광학방향과 수직일 때 최대치를 가지며, 수평일 때 최소치 0을 가지는 함수로서, 물리적 의미는 그 면에서의 반사도에 해당하게 되는데, 이는 식(2)로 나타낼 수 있다.

$$R(p,q) = 1 / \sqrt{1+p^2+q^2} \tag{식(2)}$$

여기서 p, q는 각각 z\_x(x,y), z\_y(x,y)이며, sqrt(·)는 자승근을 나타낸다. 식(1)에서 E(x,y), R을 알 때, p와 q를 구하는 편미분 방정식의 해를 구하면, 위벽면의 3차원적인 형태인 z(x,y)는 식(3)에 의해 구할 수 있다.

$$z(x,y) = z_0(x_0,y_0) + \int_c (p dx + q dy) \tag{식(3)}$$

여기서 z\_0(x\_0,y\_0)는 심도(depth)를 알고 있는 영상상의 한점을 표시하며, c는 심도를 추정하고자 하는 위치까지의 곡선을 나타낸다. 벽면의 수직벡터 n을 p, q에 의해 나타내면 식(4)와 같다.

$$n = (-p(x,y), -q(x,y), 1)^T \tag{식(4)}$$

식(4)로 표시된 n은 벽면의 x, y 방향 접선벡터 (1,0,p)^T, (0,1,q)^T와는 서로 내적이 0인 관계 즉, 수직인 관계에 있음을 알 수 있다. R(p,q)로 표시된 식(1)의 해를 구함에 있어, 벽면이 광학방향과 거의 일치하는 경우에는, p, q가 각각 무한대의 값을 가질 가능성이 있게 되며, 이를 피하기 위해서는 매핑이 필요하다. Ikeuchi와 Horn은 식(5)와 같은 매핑을 통해, (p,q)상의 모든 점을 반경이 2가 되는 원의 내부로 가계함으로써 이 문제를 해결하였다[6].

$$f = 2p / (1 + \sqrt{1+p^2+q^2}) \tag{식(5)}$$

$$g = 2q / (1 + \sqrt{1+p^2+q^2})$$

여기서 p, q는 각각 식(6)을 통해 f, g로 부터 계산할 수 있다.

$$p = 4f / (4-f^2-g^2) \tag{식(6)}$$

$$q = 4g / (4-f^2-g^2)$$

따라서 식(1)은 식(7)과 같이 표시된다.

$$E(x,y) = R(f,g) = (4-f^2-g^2) / (4+f^2+g^2) \tag{식(7)}$$

식(7)을 만족하는 면은 무수히 많아서, 그 해를 구할 수 없는 이른바, ill-posed 문제이며, 이러한 종류의 문제를 풀기 위해서는 사전에 그 해가 가지고 있는 고유의 특성을 해의 제약조건으로 하는 조절(Regulation)의 기법이 사용된다. 위벽면은 일반적인 표면과는 달리 전체적으로 매우 완만한 표면이기 때문에, 이를 제약조건으로 하는 범함수(functional)를 식(8)과 같이 정의할 수 있다.

$$J [E(x,y) - R(f,g)]^2 dx dy + \lambda \iint (f_x^2 + f_y^2 + g_x^2 + g_y^2) dx \tag{식(8)}$$

식(8)의 첫째항은 밝기오차에 해당하는 것으로 식(7)을 만족하는

경우 0이 되지만, 0이 되지 않더라도 위벽면 특성의 제약조건과의 절충을 통해 식(8)을 최소화하는 해를 구해야 한다. 둘째항의  $f_x, f_y, g_x, g_y$ 는 각각  $f, g$ 의  $x, y$  방향 편미분을 나타내며, 이들 지승합의 적분은 완만한 표면이라는 제약조건을 벗어나는 정도 즉, 급격한 변화 정도를 의미하게 된다. 식(8)을 최소화하는 해의 필요조건은 식(8)을  $f$  및  $g$ 로 편미분한 식이 0이 되는 것으로, 이러한 조건을 만족하는 식을 Euler 방정식이라 하며 식(9)로 나타낼 수 있다.

$$(E-R)R_f + \lambda \nabla^2 f = 0 \tag{식(9)}$$

$$(E-R)R_g + \lambda \nabla^2 g = 0$$

여기서  $R_f, R_g$ 는 각각 함수  $R$ 의  $f, g$ 에 대한 편미분 함수로 식(10)으로 표시할 수 있으며,  $\nabla$ 는 Laplacian operator이다.

$$R_f = -16f / (4 + f^2 + g^2)^2 \tag{식(10)}$$

$$R_g = -16g / (4 + f^2 + g^2)^2$$

식(9)를 finite difference 근사법을 이용해서, 컴퓨터에서 계산 가능한 식으로 변환하면 식(11)과 같다.

$$f_{ij} = F_{ij} + (\epsilon^2/4\lambda) (E_{ij} - R(f_{ij}, g_{ij})) R_f(f_{ij}, g_{ij}) \tag{식(11)}$$

$$g_{ij} = G_{ij} + (\epsilon^2/4\lambda) (E_{ij} - R(f_{ij}, g_{ij})) R_g(f_{ij}, g_{ij})$$

여기서  $F_{ij}, G_{ij}$ 는 각각  $f_{ij}, g_{ij}$ 에 이웃한 4좌표에서의 평균치로서 식(12)로 나타낼 수 있으며  $\epsilon$ 는 화소간의 간격을 나타낸다.

$$F_{ij} = (f_{ij+1} + f_{i+1j} + f_{ij-1} + f_{i-1j}) / 4 \tag{식(12)}$$

$$G_{ij} = (g_{ij+1} + g_{i+1j} + g_{ij-1} + g_{i-1j}) / 4$$

식(11)은 대개 반복기법을 적용해서 수렴하는 값을 그 해로 하게 되는데, 이를 계산하는 반복식은 식(13)과 같다.

$$f_{ij}^{k+1} = F_{ij}^k + (\epsilon^2/4\lambda) (E_{ij} - R(f_{ij}^k, g_{ij}^k)) R_f(f_{ij}^k, g_{ij}^k) \tag{식(13)}$$

$$g_{ij}^{k+1} = G_{ij}^k + (\epsilon^2/4\lambda) (E_{ij} - R(f_{ij}^k, g_{ij}^k)) R_g(f_{ij}^k, g_{ij}^k)$$

식(13)에서 구한  $f_{ij}, g_{ij}$ 는 식(6)을 이용해서  $p_{ij}, q_{ij}$ 로 변환할 수 있다. 여기서 최종적인 위벽면의 3차원적인 형체인  $z(x,y)$ 를 구하기 위해서는 식(3)으로 계산할 수 있으나, 실제로는 잡음이 첨가된 영상에서 계산된  $p_{ij}, q_{ij}$ 가 오차를 가지게 되므로 적분하는 과정에서 오차 누적으로 인해 직접 적용하기에는 문제가 있다. 따라서  $p_{ij}, q_{ij}$ 로부터  $z(x,y)$ 를 직접 계산하는 기법을 적용하여야 한다. 이를 위해 식(14)와 같은 범함수를 최소화하는 해를 최종해로 가정한다.

$$\iint (z_x - p)^2 + (z_y - q)^2 dx dy \tag{식(14)}$$

이 가정의 타당성은  $p, q$ 의 정의가  $z_x(x,y), z_y(x,y)$ 인 것에서 바로 증명될 수 있다. 식(15)를 만족하는 해의 필요조건은 식(15)와 같다.

$$\nabla^2 z = p_x + q_y \tag{식(15)}$$

식(15)를 근사화하고, 반복적인 기법을 적용하여 다시 정리하여 보면 식(16)과 같다.

$$z_{ij}^{k+1} = Z_{ij}^k - (\epsilon/4) (h_{ij} + v_{ij}) \tag{식(16)}$$

여기서  $Z_{ij}$ 는  $z_{ij}$ 에 이웃한 4좌표에서의 평균치로서 식(17)로 나타낼 수 있으며  $\epsilon$ 는 화소간의 간격을 나타낸다.

$$Z_{ij} = (z_{ij+1} + z_{i+1j} + z_{ij-1} + z_{i-1j}) / 4 \tag{식(17)}$$

또한  $h_{ij}, v_{ij}$ 는 각각 편미분  $p_x, q_y$ 의 추정치로서 식(18)로 나타

낼 수 있다.

$$h_{ij} = (p_{i+1j} - p_{i-1j}) / 2 \tag{식(18)}$$

$$v_{ij} = (q_{ij+1} - q_{ij-1}) / 2$$

상술한 바와 같이 본 연구에서 사용한 음영분석기법은 내시경 영상에 대한 사전정보를 바탕으로 식(8)과 식(14)를 최소화하는  $z(x,y)$ 를 반복적으로 찾아가는 알고리즘이다.

### 5. 결 과

본 연구에서는 위모형속에 전자내시경(Welch Allyn사의 Video-Endoscope™)을 삽입하여, 개인용 컴퓨터에 내장된 영상신호 디지털화 장치란 통해 영상을 획득하였다. 획득된 영상의 양식은 그 해상도가 512x512, 대조도가 256으로 Binary 화형이다. 512x512의 전체 영상 중에서 내시경영상은 중앙의 일부분만 해당하게 되는데, 본 연구에서는 중앙의 256x320을 대상으로 하였고, 이를 8x8 크기의 창내에서 평균함으로써, 32x40 크기의 영상 데이터를 전체적인 벽면의 형체를 구하는데 사용하였다. 또한 식(1)의 우변과 그 크기를 조절하기 위해서 영상상의 대조도 값을 그 최대치인 256으로 정규화하여 좌변에 해당하는  $E(x,y)$ 는 0에서 1사이의 값을 갖도록 하였다. 그림1은 디지털화된 내시경의 영상중에서 중앙의 256x320에 해당되는 것으로 위모형의 분문에 내시경 카메라의 끝부분이 있을 때 획득한 것이다. 따라서 영상의 아래 부분에 위의 소만부가 보이고 있으며, 윗부분에는 대만부가 보이고 있다.

그림2는  $\lambda$ 의 값에 따른  $f_{ij}, g_{ij}$ 의 추정에 있어서의 밝기오차, 급격한 변화정도 및 전체오차의 수렴치를 나타낸 것이다. 여기서  $\lambda$ 가 0.3일때 전체오차가 최소가 됨을 알 수 있었으며, 0.2이하에서는 오차가 급격히 증가하였다. 이는 수렴속도와도 관계되는 것으로, 수렴의 폭이 커서 속도가 빠를 경우 최적추정치를 벗어나 발산함을 의미한다.  $\lambda$ 가 1일 때는 밝기오차와 급격한 변화정도가 균형을 이루어 거의 비슷한 값을 가지게 되며, 그 이상일 경우에도 점차 발산하였다. 그림 3은  $\lambda$ 가 0.3일 경우 즉 최적의 추정치일 경우에 반복이 진행됨에 따라 오차의 변화해가는 모양을 나타낸다. 여기서 급격한 변화 정도는 15회까지는 증가하다가 그 이후에는 완만히 감소하면서 수렴하게 되는데, 밝기오차는 변화 정도가 증가하는 15회까지에서 급격히 감소하다가 그 이후에는 완만히 감소함을 알 수 있다. 따라서 15회까지에서의 전체오차는 급격한 변화정도를 증가시키는 대가로 감소하게 됨을 추정할 수 있다. 여기서 계산한  $f_{ij}, g_{ij}$ 는 식(6)에 의해  $p_{ij}, q_{ij}$ 로 변환할 수 있고, 이를 이용해서 식(16)의 반복식으로  $z(x,y)$ 를 계산할 수 있다. 그림4는 최종적으로 계산한  $z(x,y)$ 의 형태를 3차원적으로 나타낸 것으로  $z$ 축 상에서 큰 값을 가질 수록 면의 위치가 카메라와 가까움을 나타낸다. 벽면상에서 오탁한 부분이 카메라에서 보았을 때 가장 멀리 보이는 위모형의 대만부에 해당한다.

6. 결론

본 연구에서는 전자내시경 영상의 3차원 재구성을 위해, 우선 내시경의 기기적인 특성에 의해 내시경 영상에 어떠한 고유의 특성이 있게 되는지를 분석하고, 진단 대상이 되는 위벽면 자체의 특성으로 인한 내시경 영상의 특성을 분석하였다. 이를 기반으로 하여 일반적으로 3차원 정보의 추출에 사용되는 알고리즘들을 분석하여 내시경 영상에의 적용 여부를 평가하였다. 이 과정에서 일반적인 3차원 정보추출 기법 중에서 음영분석기법이 가장 적합한 것으로 밝혀졌다. 따라서 타 기법과의 병행처리를 시도할 경우, 이를 기준으로 해서 다른 결과를 병합하여야 할 것임을 알 수 있다. 알고리즘화를 위해 Horn의 변화법을 도입하여 반복적으로 구현하였고, 이의 타당성 여부를 증명하기 위하여 위모형에서 상용 내시경을 사용하여 영상을 획득하고, 이 중에서 1220(=32x40)개의 데이터클 샘플링하여 이를 기준으로 시뮬레이션을 수행하였다. 이 때 Lagrangian multiplier  $\lambda$ 의 값이 0.3인때, 안고리즘의 수렴성 및 전체오차 특성이 가장 좋았으며, 가장 정확한 위벽면의 재구성이 이루어짐을 밝혔다. 본 연구는 국내의 수 많은 소화기 관련 환자류 진단하는데 있어, 병원에서 시행하는 내시경 검사의 결과를 3차원적으로 분석하고자 하는 것이다. 물론 임상 의사는 순차적인 영상을 관찰함으로써, 머리 속으로 그 모양을 추정하겠지만 이를 컴퓨터를 이용하여 보다 객관화, 정량화한 영상을 재구성하는 것은 보다 정확한 진단에 큰 도움을 줄 것이다. 또한 3차원적인 정보를 가지고 있는 면은 내시경에서 위 내의 어느 방향으로도 그 2차원적인 영상을 재구성할 수 있기 때문에, 임상 의사가 관심을 두지 않아 관찰하지 못한 병변에 대한 정확한 정보를 검사 후에 컴퓨터를 통한 분석을 통해 알 수 있을 것으로 판단된다.



Figure 1. Stomach surface model image (256x256) acquired by videoendoscope

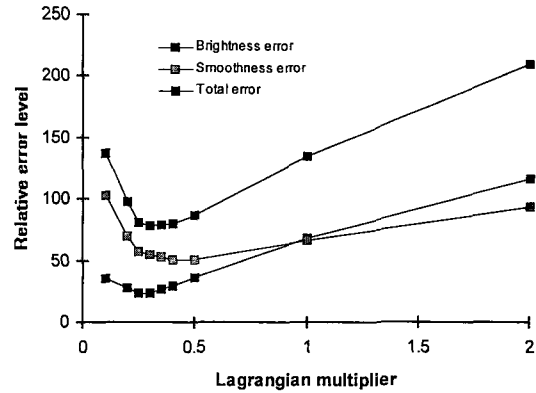


Figure 2. Convergence levels of brightness, smoothness, and total error for various Lagrangian multipliers

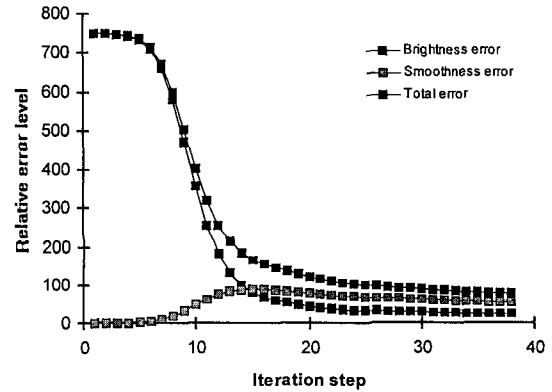


Figure 3. Convergence curves of brightness, and smoothness, and total errors( $\lambda=0.3$ )

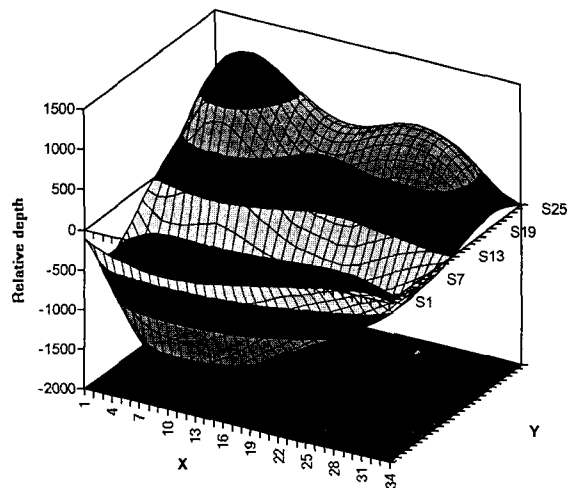


Figure 4. 3-D graph of estimated stomach surface model