

セルアニメーションからの三次元キャラクタの生成

Generating 3-D Character Model from Cel Animation

橋本 泰成 小野 豊 橋本 良太 五十嵐 健夫 西田 友是
Yasunari Hashimoto Yutaka Ono Ryota Hashimoto Takeo Igarashi Tomoyuki Nishita

東京大学
The University of Tokyo

概要

本研究ではキャラクタの周囲をカメラが回転する二次元のセルアニメーションからキャラクタの三次元形状を自動的に生成する手法を提案する。一般に、セルアニメーションのキャラクタは三次元的に正確な形状を意識して描かれているわけではないので、画像列から適切な三次元形状を定義するのは困難である。そこで、本研究では対象キャラクタが手足などのコンポーネントからできており、それぞれが丸みを帯びていると仮定し、胴体となる形状に対してそれらのコンポーネントを追加していくことで構造的に妥当な三次元形状を生成する。我々は実験用のセルアニメーションに対し提案法を適用し、その有効性を検討した。

1 はじめに

コンピュータグラフィクスやコンピュータビジョンの分野において、複数の二次元画像からの三次元モデルの生成は重要なトピックの一つであり多くの研究がなされている。しかし、それらの研究の多くは人間や動物など正確な三次元形状が存在するオブジェクトに対して行われており、セルアニメーションのキャラクタのように明確な三次元形状が定義されていないものから三次元モデルを自動的に生成する手法は我々の知る限りではまだない。そこで、本論文ではセルアニメーションから三次元キャラクタを自動的に生成する手法を提案する。提案手法は、セルアニメーションへの影付け [4] や簡易なモデリングシステム、ゲームやエンターテイメントなどへの応用が考えられる。

一般に、視点位置が与えられていない場合に複数の画像から物体の三次元形状を生成する際は、各画像間での点や線の対応を求め、そこからキャラクタの移動方向やカメラの位置を推定する必要がある。コンピュータビジョンの手法では、オプティカルフロー [3] や、特徴点の追跡を用いた手法 [5] 等が広く用いられているが、フレームレートが小さく、輝度値が離散的に変化するセル

アニメーションへそれらの手法を適用することは困難である。そこで、提案法ではキャラクタの輪郭形状のみを用いて三次元形状を生成する。



図 1: 入力セルアニメーションの数フレームの画像

本研究はまだ開発段階にあり、入力アニメーションを、図 1 のペンギンのような、キャラクタが手足などを移動させておらず、カメラがキャラクタを中心とする垂直な軸で回転しているものに限定する。以下ではそのような入力アニメーションを用いての実験結果を示し、提案手法の適用可能性を議論する。

2 関連研究

複数の輪郭形状のみを用いて三次元モデルを生成する代表的な手法として、visual hull [2] がある。これは輪郭と、カメラ位置により形成される錐体の積をとることで、対象物体が存在しうる空間を削り落としていく手法である。この手法は正確な三次元形状をもつ物体に対しては有効であるが、セルアニメーションのキャラクタの場合、アニメータが物理的な正確さよりも見た目の美しさを優先したためにキャラクタが手足等を動かしていないにもかかわらず見る角度によって三次元形状としての不整合（以降、このことを「視点依存のゆがみ」と呼ぶことにする）が生じることが多くある。そのためセルアニメーションのキャラクタについては全ての輪郭を保存するような visual hull を定義するのは一般に不可能である。そのうえ、場合によっては過剰な錐体間の積演算によって手や足などのキャラクタの重要な特徴が失われてしまう可能性がある。

一方、キャラクタの形状が丸みを帯びているという仮定の下で、少数の輪郭から三次元モデルを生成する

Teddy[1] というスケッチシステムが五十嵐らにより提案されている。セルアニメーションのキャラクタは一般に丸みを帯びているものが多く、そのようなキャラクタの簡単な三次元形状を作るのにこの手法は有効である。

上で述べたように、様々な視点から見た輪郭を全て保存して三次元のセルアニメーションのキャラクタを作るのは一般に不可能である。そこで、提案手法ではキャラクタが胴体や手足でそれぞれ丸みを帯びているという仮定の下で [1] の手法の中の 2 つの操作、すなわち (a) 閉曲線から丸みを帯びた三次元形状を生成する操作と (b) 二次元のストロークをもとに三次元形状に突起を追加する操作を用いて目的の三次元形状を生成する。提案法ではキャラクタの胴体部分を (a) の操作で、手足などの構成要素を (b) の操作を用いて生成する。最終的な形状は、入力アニメーションのキャラクタと手足などが同じ構成要素からなるという意味で、構造的に正しいものとなる。

3 提案手法の概要

提案手法は、入力のセルアニメーション¹から動物のようなキャラクタの三次元形状を生成するが、それらは以下のような仮定を満たしているものとする。まず、対象となるキャラクタは、胴体とそれに接続する手足等の特徴的な凸部分の集合よりなる。胴体を含むそれらの凸部分を以下ではコンポーネントと呼ぶ。各コンポーネントは丸みを帯びた形状を持つ。次に、入力のセルアニメーションは、正射影のカメラによる連続した画像列であり、対象キャラクタは動かず、カメラ位置のみが変化する。ただし、キャラクタがセルアニメーションに特有の、視点依存のゆがみを生じることは許すものとする。さらに、図 1 のペンギンの手足やくちばしのように、アニメーション中で各コンポーネントの長径が一度以上画像平面の上ののりとする。これらの仮定の下でも生成できる三次元形状にはあいまいさが残るが、本手法ではその中から可能な三次元形状の一つを生成する。

三次元モデルの生成アルゴリズムは以下のものである。まず、各画像の輪郭からキャラクタのコンポーネントを推定し (第 4 節)、フレーム間でそれらの対応関係を求める (第 5 節)。そして、各コンポーネントの対応関係からコンポーネント間の相対位置を推定し、最後に、各コンポーネントについて、その形状を最もよくとらえている輪郭を用いて、五十嵐らの手法 [1] により胴体に次々とコンポーネントを追加する (第 6 節)。

4 輪郭を用いたコンポーネント推定

セルアニメーションのキャラクタの三次元形状生成にあたり、まずキャラクタを手足などのコンポーネントに

切り分ける。コンポーネントは三次元的に凸となる部位であるが、その二次元への投影も凸構造となるので、輪郭をもとに二次元でコンポーネントの推定を行う。輪郭形状をコンポーネントに切り分ける研究としてはスケルトンを利用した手法 [6] などがすでに存在するが、提案法では輪郭の曲率を用いたアプローチでコンポーネントを推定している²。なお、以下では文脈上明らかである場合には、三次元上で凸となる部分に加え、二次上で凸となる部分も区別無く単にコンポーネントと呼ぶ。輪郭は点列からなる閉曲線とし、輪郭のコンポーネント分けは以下のアルゴリズムを各フレームの画像に対して適用することで行う。

1. 輪郭上の各点の曲率から適当な閾値を用いて特徴となる点 (以下、特徴点と呼ぶ) を求め (図 2(a))、それぞれの特徴点を直線で結ぶことで多角形を形成する (図 2(b))
2. 多角形上の頂点の中で凹となっている点を新たな特徴点とする (図 2(c))
3. 隣接した特徴点同士を直線で結ぶことで、直線と外側の輪郭による凸多角形を形成する (図 2(d))
4. 3 の凸多角形について、直線部分の長さを l_a とし、輪郭上で直線から最も遠い点から直線までの長さを l_b とおき、 $l_a < l_b$ を満たす場合にその多角形をコンポーネントとみなす
この判定は胴体の一部等を誤ってコンポーネントとみなさないために行っている (図 2(e) で横線部分はコンポーネントとみなすが、縦線部分は条件を満たさないでコンポーネントとみなされない)
5. 輪郭からコンポーネントを除き、残された特徴点で多角形を形成する
6. 2,3,4,5 の処理を多角形の各点の内角が全て 180 度以下になるまで繰り返し、残った多角形の内部をキャラクタの胴体とする (図 2(f) の白い部分が胴体となる)

このアルゴリズムは、キャラクタの形状が単純で凹凸がはっきりした形状をしている場合、意図したようにいく場合が多いが、コンポーネントが複雑に重なり合う場合などに、誤認識してしまう場合もある。

²冒頭でも述べたように本研究はまだ開発段階であり、このアプローチが最適であると主張するものではない。[6] やその他の手法を用いた場合の実験も今後の課題である。ただし、我々の行った実験ではスケルトンを用いた手法は輪郭の細かいノイズに左右されやすく、必ずしもよい結果を出さなかったことを述べておく。

¹実験の際は、簡単のためあらかじめ背景部分を塗りつぶしている

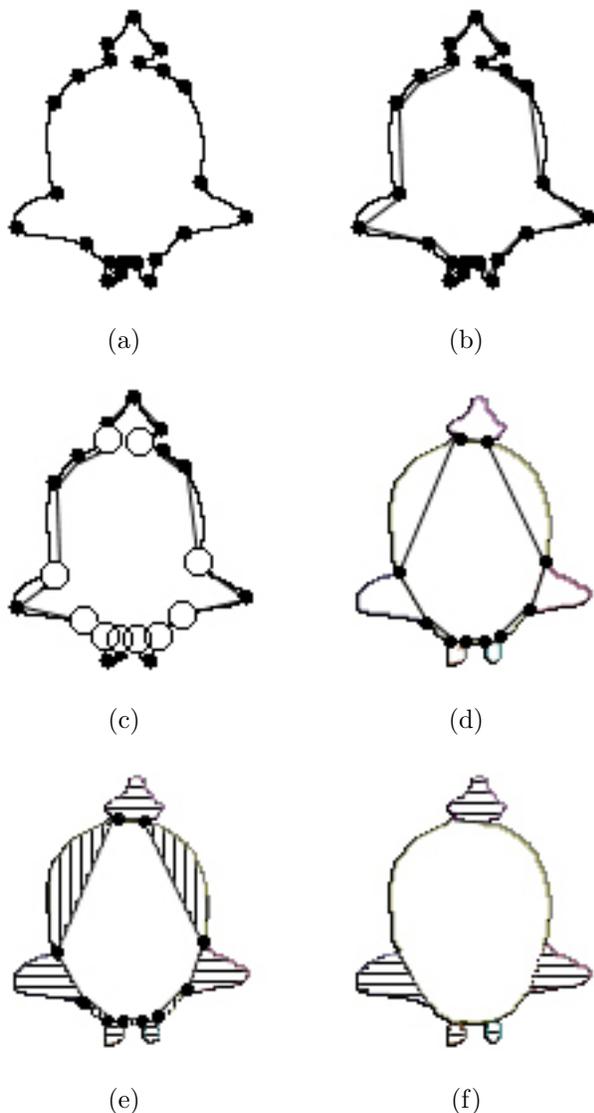


図 2: 輪郭を用いたコンポーネント化の過程。図の見方は本文を参照のこと

5 フレーム間でのコンポーネントの対応

次に各フレームの間でそれぞれのコンポーネントについて、対応関係を以下の手法で求める。

1. 隣接するフレーム間での対応づけ

入力セルアニメーションはフレーム間での移動量が小さいことを仮定しているので、 i 番目のフレームと $i+1$ 番目のフレームで位置が近い³コンポーネント同士を対応しているコンポーネントとし、 i 番目のフレームのどのコンポーネントにも対応づけられないコンポーネントはそのフレームから新しく現れたコンポーネントとする。逆に、 i 番目のフレームで存在したコンポーネントが $i+1$ 番目のフレーム

³重心の位置の近さと面積の重なり具合を判断基準としている

ではどのコンポーネントとも対応付けられなくなることがある。そのようなコンポーネントは $i+1$ フレームで消失したと考える。

2. ノイズの除去

先に述べたように、コンポーネント分けの際にノイズが入る場合があるので、1~2フレームだけ現れるようなコンポーネントは無視することでノイズを除去する。

3. 離れたフレーム間での対応づけ

胴体に隠れた手が再び出てくるなど、離れたフレーム間でもコンポーネントが対応付けられることがある。そこで、コンポーネントに対し出現期間という概念を導入する。コンポーネントの出現期間はそのコンポーネントに対応付けられるコンポーネントが存在しているようなフレームの集合として定義され、出現期間の重ならないコンポーネント同士についてはコンポーネントの大きさや消失時、出現時の移動方向ベクトルなどを用いて、対応がとれるかどうかを判断する。そして対応がとれると判断できるものについてはそれらを同一のコンポーネントであるとみなし、お互いを対応付ける。図 3 はその例で、図 3 (a) で灰色で示された部分のコンポーネントは移動ベクトルと大きさから判断して回転の途中で一度、胴体に隠れて (図 3(b))、その後のフレームで出現した同一のコンポーネント (図 3(c)) として対応付けられる。

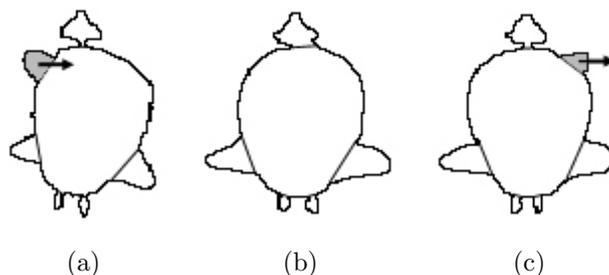


図 3: 離れたフレーム間でのコンポーネントの対応。(a) で灰色で示されたコンポーネントは移動ベクトルから (b) で一度胴体に隠れ、(c) で再び現れたものとする。矢印は移動ベクトル

6 三次元形状の生成

前節までの処理でキャラクタをコンポーネントに分割し、フレーム間での対応をとることができた。これらのコンポーネントには三次元的な情報が欠如しているため、それらを推定する方法を以下で説明する。

先にも述べたとおり、セルアニメーションのキャラクタ形状は視点に依存したゆがみをもっているため全ての

輪郭を保存した三次元モデルを生成するのは一般に不可能である。そこで、提案法では各コンポーネントについて代表的な輪郭をひとつ選びその輪郭を保存するように該当コンポーネントの三次元形状を生成する。輪郭から三次元形状を生成する際には、コンポーネントが丸みを帯びているという仮定をもとに Teddy [1] の手法における (a) 閉曲線から丸みを帯びた三次元形状を生成する操作と (b) 二次元のストロークをもとに三次元形状に突起を追加する操作を用いている。

アルゴリズムの詳細を述べる前に、必要な用語の定義をしておく。「手」「足」「胴体」など、切り分けられたそれぞれの形状に対してそれに関連付けられる複数の輪郭が存在するが、この節では前者のことをコンポーネントと呼び、後者については単に輪郭と呼ぶことにする。また、 i 番目のコンポーネントを $component(i)$ 、 t 番目のフレーム画像において $component(i)$ に対応付けられる輪郭を $silhouette(i, t)$ と表記することにする。

具体的な処理の流れは以下ようになる。まず、各 i に対し、 $silhouette(i, t)$ の面積が最大になるような t を求め、 t_i とおく。そして、 $silhouette(i, t_i)$ を、 $component(i)$ を代表する輪郭 $silhouette(i)$ とおく。次に、[1] の手法を用い、胴体コンポーネントの代表輪郭を三次元化し、初期形状とする。三次元座標の原点を初期形状の重心におき、以後は [1] の手法を用いて初期形状に突起状の形状を付け加えていく形で残りのコンポーネントを順次追加していき、最終的な形状を得る。

三次元化したコンポーネントを胴体の三次元形状に追加する際、コンポーネントの三次元位置（この場合、コンポーネントの代表輪郭の含まれるフレームでのカメラ位置）を求める必要がある。提案法では、追加しようとしているコンポーネントの三次元位置を、他のコンポーネントの三次元位置と二次元的な見えの情報を元にして決定する。例として、 $component(i)$ を元にして $component(j)$ の三次元位置を求める場合を考える（フレーム t_i におけるカメラ位置はすでに求められているものとする）。このとき、 $component(j)$ の三次元位置

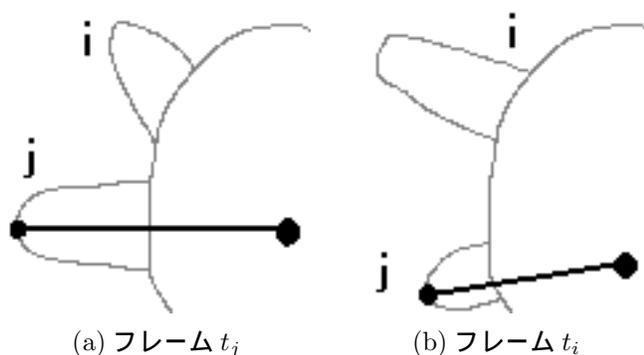


図 4: コンポーネントの三次元的な位置の決定方法。胴体コンポーネントに i と j の 2 つのコンポーネントが付属している。

を求めるためにフレーム t_j と t_i を取り出し 2 つの輪郭 $silhouette(j, t_j)$ と $silhouette(j, t_i)$ を参照することになる (図 4)。そして、図 4 のように胴体コンポーネントの輪郭の重心と $component(j)$ の輪郭の先端を結ぶ線分を考える。 t_j の選び方と入力アニメーションの仮定から、 $component(j)$ から伸びる線分の長さはフレーム t_j において最大となることが期待される。そこでこの線分の長さ l_j を $component(j)$ の長さとおく。また、フレーム t_j と t_i における $component(j)$ の先端の座標をそれぞれ (x_{j,t_j}, y_{j,t_j}) と (x_{j,t_i}, y_{j,t_i}) とおくと、正射投影の仮定よりこれらの三次元座標はそれぞれ $(x_{j,t_j}, y_{j,t_j}, z_{j,t_j})$ 、 $(x_{j,t_i}, y_{j,t_i}, z_{j,t_i})$ で表せる。この時、フレーム t_j から t_i にかけてカメラ位置が X 軸を中心に θ 、Y 軸を中心に ϕ だけ回転したと考え、以下の式を解くことで回転量の候補を求めることができる。

$$\begin{aligned} l_j &= \sqrt{x_{j,t_j}^2 + y_{j,t_j}^2} \\ &= \sqrt{x_{j,t_j}^2 + y_{j,t_j}^2 + z_{j,t_j}^2} \\ &= \sqrt{x_{j,t_i}^2 + y_{j,t_i}^2 + z_{j,t_i}^2} \end{aligned} \quad (1)$$

$$\begin{pmatrix} x_{j,t_i} \\ y_{j,t_i} \\ z_{j,t_i} \end{pmatrix} = M_2(\phi)M_1(\theta) \begin{pmatrix} x_{j,t_j} \\ y_{j,t_j} \\ z_{j,t_j} \end{pmatrix} \quad (2)$$

ただし、

$$M_1(\theta) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix} \quad (3)$$

$$M_2(\phi) = \begin{pmatrix} \cos \phi & 0 & \sin \phi \\ 0 & 1 & 0 \\ -\sin \phi & 0 & \cos \phi \end{pmatrix} \quad (4)$$

フレーム t_i でのカメラ位置は既に計算されているので、 θ と ϕ の値からフレーム t_j でのカメラ位置を求めることができる。

また、初期形状に最初のコンポーネントを追加するときには位置を参照できるコンポーネントが存在しないが、この場合は胴体コンポーネントの三次元位置を基準としてカメラ位置を推定する。

このようにして求められたカメラの移動には曖昧さが残り、複数の可能性が存在するがこの曖昧さは複数のコンポーネントとの相対位置を考慮し、互いに矛盾のないカメラ移動を採用することである程度解決できる。参照できるコンポーネントの数が少ない場合等、曖昧さが完全に解決できない場合には提案法ではランダムにひとつの可能性を選択することで対処している。

7 実験結果

実験用のセルアニメーションに対し提案法を適用した例を以下に示す。図 5 は図 1 に示したセルアニメーション

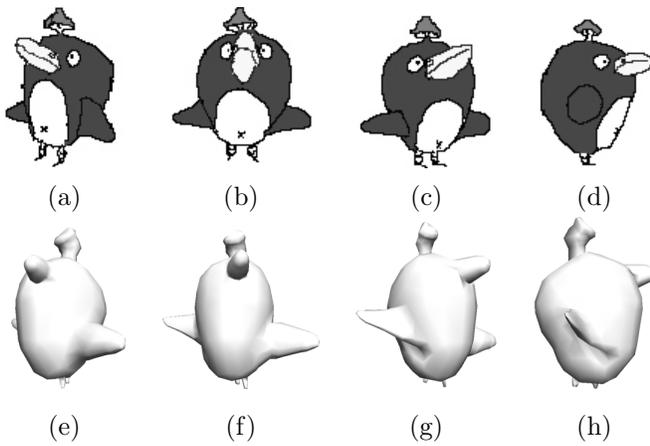


図 5: 提案法の適用結果。(a)~(d): 入力アニメーションからの画像。(e)~(h): 出力モデルを対応する視点から見たもの。アニメーションにおいて顕著な凸部分をコンポーネントとして検出できていることがわかる。

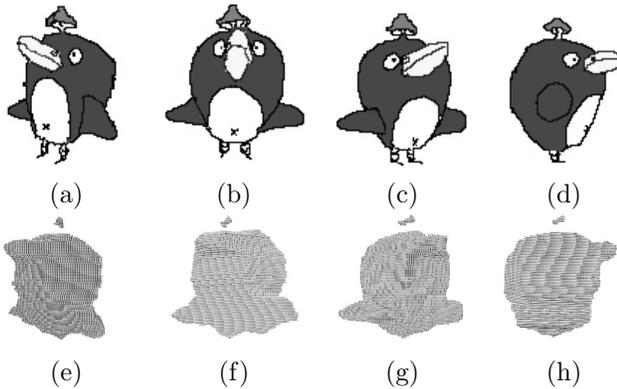


図 6: visual hull を用いて三次元モデルを生成した例。(a)~(d): 入力アニメーションからの画像。(e)~(h): 対応する視点での visual hull。キャラクター頭部のキノコが胴体から離れ、足が失われていることがわかる。

ンを入力として与えた結果である。入力フレーム数は13である。この例では、キャラクターから胴体と左右の手足、頭上のキノコ、くちばしの計7つのコンポーネントが検出された。計算時間は、コンポーネント推定、コンポーネントの対応付け、三次元形状の生成をすべてあわせて、Pentium4 1.7GHz 搭載の PC で2~3秒程度であった。

比較のために、同じアニメーションを入力として visual hull を用いて得られた結果を図6に示す。この例の場合、セルアニメーション中でキャラクターは明示的な動きを見せていないが、アニメータによる視点に依存するゆがみのために、全ての輪郭を保存するような visual hull を作ることはできなかった。その上、頭上のキノコが胴体から切り離され、足もなくなってしまっている。

提案法により得られた三次元モデルは、入力アニメーションの画像と同一の方向から見た場合でも形状が画像と異なることがあるが、コンポーネント数や相対的な位置は、おおまかではあるものの、もとの入力アニメーション

と一貫性が取れているものになっている。

8 おわりに

本論文で我々は、セルアニメーションからキャラクターの三次元形状を生成する手法を提案した。提案したプロトタイプシステムにより生成される三次元モデルは、入力アニメーション上でのキャラクターと同一の構造を持ち、従来法では困難であった視点依存のゆがみをもつキャラクターにも適用することが可能である。

本研究はまだ開発段階にあり、一つの例に対して結果を確認したのみである。今後はより良いモデルを生成できるように改良し、適用可能な範囲をより一般に拡張する必要がある。具体的な改良点としては、輪郭のコンポーネントへの分割や対応において複数の候補がある際に誤った認識をする場合があるので時系列による履歴的な閾値を用いて、それらのノイズを減らすことや、コンポーネントの三次元化に際してより詳細な検討を行うことなどがあげられる。また、本手法においている仮定を満たすセルアニメーションは非常に限られているので、今後はそれらを緩和し、動いているキャラクターや回転以外のカメラの動きなどにも対応できるようにしていきたいと考えている。

参考文献

- [1] T. Igarashi, S. Matsuoka, H. Tanaka, "Teddy: A Sketching Interface for 3D Freeform Design," Proc. SIGGRAPH 99, 1999, pp. 409-416.
- [2] A. Laurentini, "The Visual Hull Concept for Silhouette-based Image Understanding," Trans. Pattern Analysis and Machine Intelligence, 16(2), 1994, pp. 150-162.
- [3] B. Lucas, T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," Proc. 7th International Joint Conference on Artificial Intelligence, 1981, pp. 674-679.
- [4] L. Petrovic, B. Fujito, L. Williams, A. Finkelstein, "Shadows for Cel Animation," Proc. SIGGRAPH 2000, 2000, pp. 511-516.
- [5] J. Shi, C. Tomasi, "Good Features to Track," Proc. Conference on Computer Vision and Pattern Recognition, 1994, pp. 593-600.
- [6] S. C. Zhu, A. L. Yuille, "FORMS: A Flexible Object Recognition and Modeling System", International Journal of Computer Vision, 20, 1996, pp. 187-212.