

# Bölüt ve Kontur Özniteliklerini Kullanarak İmgelerdeki İnsan Hareketlerini Tanıma Recognizing Human Actions in Images Using Segment and Contour Features

Gökhan Tanışık  
Aselsan A.Ş.  
Akyurt, Ankara, Türkiye  
Email: gtanisik@aselsan.com.tr

Oğuzhan Güçlü  
Bilgisayar Mühendisliği Bölümü  
Hacettepe Üniversitesi  
Ankara, Türkiye  
Email: oguzhanguclu@cs.hacettepe.edu.tr

Nazlı İkizler-Cinbis  
Bilgisayar Mühendisliği Bölümü  
Hacettepe Üniversitesi  
Ankara, Türkiye  
Email: nazli@cs.hacettepe.edu.tr

**Özetçe** —Bu çalışmada, resimlerden insan hareketi tanıma hedeflenmiştir. Diğer çalışmalardan farklı olarak, buradaki amaç kontur tespiti ve resim bölütleme algoritmaları kullanılarak resim içerisindeki kontur ve bölütlerin insan pozunu temsil edip etmediğini araştırmaktır. Önerilen yöntemde, ilk olarak resimdeki kontur ve bölütler bulunur. Bulunan kontur ve bölütler niteliklerin elde edilmesi için kullanılır. Nitelikler çıkarıldıktan sonra, kontur ve bölütlere ayrıştırılmış olan resimler gridlere ayrılarak uzlamsal bilgi elde edilir. Her bir griddeki öznitelikler birleştirilerek bir histogram oluşturulur. Elde edilen bu histogram hareketlerin sınıflandırılması için kullanılır. Sınıflandırma yöntemi olarak Destek Vektör Makineleri (DVM) yöntemi kullanılmıştır. Elde edilen sonuçlar resimlerdeki kontur ve bölütlerin bir arada kullanmanın insan hareketini temsil etmek için daha etkili olduğunu göstermektedir.

**Anahtar Kelimeler**—Hareket Tanıma; Kontur Tespiti; Resim Bölütleme.

**Abstract**—In this paper, the aim is to recognize human actions from still images. Different from other works, the aim is to use a contour detection and an image segmentation algorithm together to see if the contours and segments are able to represent the human pose in a still image. As a solution, firstly, contours and segments are found in the image. The segments and contours are then used to extract the features. After extracting the features, the contoured and segmented images are split as a grid to get the spatial information. Then, the features in every grid are concatenated as a histogram and the histogram is used for classification. Support Vector Machines (SVM) algorithm is used for classification. The results demonstrate that using the contours and segments together in the image is a better representation for recognizing human actions in still images.

**Keywords**—Action Recognition, Contour Detection, Image Segmentation.

## I. GİRİŞ

İnsan hareketlerinin tanınması, bilgisayarlı görü literatüründe son yıllarda oldukça popüler olmuş bir konudur. Bunun nedeni, temel olarak video ve imgelerde, insan öğesinin



Şekil 1. Resimlerde insan hareketleri için örnekler.

merak edilen bir unsur olması, ve hareketlerin otomatikman tanınmasının, güvenlikten sağlık hizmetlerine kadar pekçok alanda kullanımı olmasıdır. Videolarda insan hareketini tanıma, son senelerde, nispeten oldukça fazla çalışılmış bir konudur ([10]). Videolara oranla, imgelerde varolan insan hareketini tanıma ise oldukça az çalışılmış bir konudur. Oysaki, resmin görsel bütünü otomatik olarak tanımlayabilmek için, insan hareketlerini tanımak oldukça önem taşımaktadır. Özellikle haber görsellerinde, tek bir fotoğraf ile insan hareketini anlatılmak istenmektedir ve tek bir karenin bu konuda yeterli olduğu görülmektedir (Bakınız Şekil 1). Bu durumda, otomatik analiz için, tek bir imgeden hareketin algılanması gerekmektedir.

Bu çalışmada, sabit imgelerdeki insan hareketlerinin analizi ve tanınması için bir yöntem önermekteyiz. Bu problem, arkaplan çıkarımı olmadığı için, videolara göre nispeten daha zor olabileceği bir problemdir. Bu çalışmamızda, ön plan kişi bilgisine ulaşmak için, kontur ve bölüt bilgisinden faydalanıyoruz. Kullanılan yaklaşım ise resimlerdeki kontur ve bölütlerin birlikte kullanılması ve bu ikisinin insan hareketini tanımak için yeterince bilgi sağlayıp sağlamadığının test edilmesidir. Bu çalışmada, ara bir insan pozunu belirleyicisi kullanmak yerine, sadece kontur ve segmentler kullanılmıştır. Bu durum problemi daha da zorlaştırmıştır. Çünkü arka planda bulunan dağınıklık (*clutter*) sebebiyle çok sayıda ilgisiz kontur ve bölüt elde edilmektedir. Bunlar ise insan pozunun temsilini güçleştirmektedir. Güçlüklere rağmen, bu çalışmada resimlerdeki insan hareketlerini tanımak için henüz başlangıç aşamasında olan bir algoritma sunulmuştur.

Önerilen yöntemde, kontur ve bölütleri elde etmek için literatürdeki en başarılı algoritmalarından birisi olarak bilinen Arbelaez ve diğerleri tarafından önerilen, kontur ve bölütleri

birlikte bulan teknik kullanılmıştır [9]. Resimlere ait kontur ve bölütler ayrı ayrı gridler şeklinde bölünerek uzlamsal bilgi de eklenmiş olur. Hem kontur hem de bölütler için gridlerin alt parçalarının yönelimlerinden  $180^\circ$ 'lik düzlemde  $15^\circ$ 'lik açığa sahip çözünürlükte olan histogramlar oluşturulmuştur. Her ikisi için elde edilen uzamsal histogramlar birleştirilerek,  $108 (9 \times [180/15])$  alt birimden oluşan birer histogram oluşturulur. Daha sonra kontur ve bölüt özniteliklerinden oluşan iki histogram da birleştirilir ve tek bir histogram elde edilir. Böylece sınıflandırmada kullanılacak olan öznitelik kümesi elde edilmiş olur.

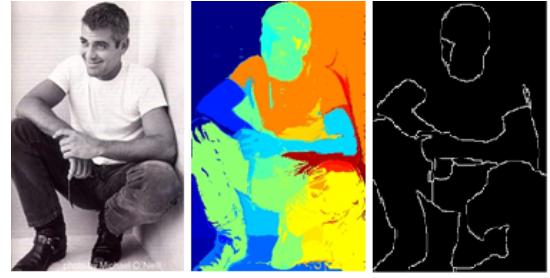
Sınıflandırma için Destek Vektör Makinesi (DVM) yöntemi kullanılmıştır ve deney sonuçları Çapraz Geçerleme (*Cross Validation*) ile elde edilmiştir. Sonuçlar, kontur ve bölüt özniteliklerinin birarada kullanılmasının sabit imgelerdeki insan hareketlerini tanıma daha etkili olduğunu göstermektedir.

## II. İLGİLİ ÇALIŞMALAR

Resimlerden insan hareketi tanıma üzerine literatürde fazla çalışma mevcut değildir. Mevcut çalışmaların ise büyük çoğunluğu videolar üzerinde gerçekleştirilmiştir ([10]). Bunun sebebi, resimler üzerinde çalışmanın videolara kıyasla daha zor bir problem olmasıdır. Sabit imgelede hareket tanımaya yönelik ilk çalışmalardan birisi üst vücut pozlarının tanınması üzerinedir [5]. Bu çalışmada insan vücuduna ait yerel niteliklerin elde edilmesi için matrislerin negatif olmayan çarpanlarına ayrılması (*non-Negative Matrix Factorization - nNMF*) yöntemi kullanılmış ve sistem etiketli bir resim veri kümesi ile eğitilmiştir. Bu çalışmanın en önemli avantajı arka plan dağınıklığından etkilenmemesidir.

İnsan pozu tespiti için nNMF yöntemini kullanan diğer bir yaklaşım ise Thureau ve diğerleri tarafından önerilmiştir [2]. Bu yaklaşım ise nitelik çıkarımı için videoları kullanmakta fakat hareketleri sınıflandırmak için resim ya da video kullanabilmektedir. Elde edilen niteliklerin indirgenmesi için Histogram of Gradients (HoG) yöntemi kullanılmıştır. Hareketlerin sınıflandırılması için histogram eşleştirme metodu tercih edilmiştir. Diğer bir yöntem ise Yang ve diğerleri tarafından önerilmiştir [3]. Bu yöntemde insan pozu tanımının yanında gizli pozların (golf oynamak gibi) bulunması da amaçlanmıştır. İnsan vücudunun uzuvları tespit edilip *poseletler* olarak kullanılmıştır. Benzer şekilde HoG yöntemi kullanılarak nitelikler elde edilmiş ve DVM algoritması kullanılarak sınıflandırma gerçekleştirilmiştir.

Öte yandan, Eicher ve diğerleri probleme yeni bir yaklaşım getirmişlerdir [4]. Çalışma, resimlerdeki insanların kafa, gövde, kol pozisyonları gibi uzlamsal yerleşimlerinin belirlenmesini hedeflemektedir. Bu bildiredeki çalışmaya da ilham kaynağı olan yöntem ise İkizler ve diğerleri tarafından geliştirilmiştir [1]. Burada ise, insan silüetinin elde edilmesi için bir insan ve poz tespiti yöntemi kullanılmıştır. Daha sonra silüetin parçaları dikkörtgensel bölgelere ayrılmıştır. Bu dikkörtgenlerin yönelimleri kullanılarak bir öznitelik histogramı elde edilmiştir. Bizim çalışmamızda, ek olarak ara bir insan pozu belirleyicisi kullanmadan, sadece bölüt ve kontur bilgileri kullanarak insan hareketini tanımak için bir yöntem geliştirilmiştir.



Şekil 2. Kontur tespiti ve bölütleme örneği. Soldan sağa doğru sırasıyla; orijinal, bölütlenmiş ve konturlarına ayrılmış resimler gösterilmiştir.

## III. ÖNERİLEN YÖNTEM

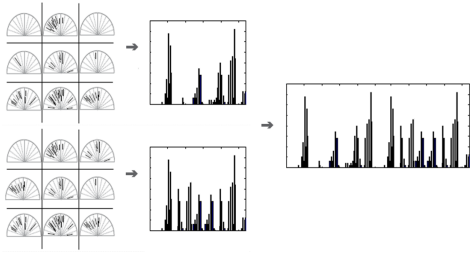
Bu çalışmadaki amaç insan hareketini tanıma için sadece kontur ve bölütlerin yeterli derecede bilgi içerip içermediğinin incelenmesidir. Kontur tespiti bir resimdeki bölgelere ait sınırların elde edilmesi işlemidir. Bu alan üzerine son yıllarda yoğun çalışmalar yapılmaktadır. Örneğin, Kass ve diğerleri tarafından önerilen yöntem bu alanda temel oluşturmuştur ve sonrasında bu konu üzerine yapılan araştırmalar önem kazanmıştır [8].

Resim bölütleme, resimlerin daha anlamlı ve analiz edilmesi daha kolay piksel kümelerine (*superpixels*) ayrılması işlemidir. Graph Based [6] ve N-Cuts [7] bölütleme algoritmaları bu alanda iyi bilinen bölütleme yöntemleridir.

Bu çalışmada önerilen yaklaşımda resimlerin kontur ve bölütlerini birlikte veren Arbelaez ve diğerleri tarafından geliştirilen algoritma kullanılmıştır [9]. Bu algoritmada resim bölütleme problemi kontur tespitine indirgenmiştir. Kontur tespiti için resme ait parlaklık, renk ve doku gibi yerel nitelikler spektral kümeleme kullanılarak birleştirilmiş ve genelleştirilmiştir. Bölütlerin elde edilmesi için ise hiyerarşik bölge ağacı kullanılmıştır.

Çalışmanın temelini resimlerdeki kontur ve bölütler oluşturmaktadır. Şekil 2 örnek bir resme ait elde edilen kontur tespiti ve bölütleme sonucunu göstermektedir.

Çalışma gri-ölçekli resimler üzerinde işlem yapıldığından, varsa RGB olan resimler öncelikle gri-ölçekli hale dönüştürülür. Sonrasında, resme ait kontur ve bölütler çıkarılır. Bölütleme algoritması resimdeki her bir bölüte bir etiket atar ve bölütün içindeki piksel değerlerini o etiket değeriyle değiştirerek etiketlenmiş bir resim döndürür. Resimdeki her bölüt ayrı olarak işlenir; bu amaçla her birisi için bir siyah-beyaz resim oluşturulur. Bu resimde bölüte ait pikseller 1 (beyaz), diğer kısımlar ise 0 (siyah) olarak değiştirilir. Kontur tespiti sonucunda elde edilen resim zaten siyah-beyazdır. Bu siyah-beyaz resimler eşit boyutta parçalardan oluşan  $N \times N$ 'lik gridlere bölünür. En iyi sonuçlar  $N = 3$  olduğu durumda elde edildiği için  $3 \times 3$ 'lük gridler kullanılmıştır. Bu bölme işlemi sonunda, gridin herhangi bir parçasında bir bölüte ya da kontura ait birden fazla kesim bulunabilir. Dolayısıyla, parçanın içerisinde yer alan bu kesimler birbirinden ayrı bölütler ya da konturlar olarak ele alınırlar. Farklı bir deyişle, bir bölüt 9 kesime ayrılabilir, buna ek olarak da her bir grid parçası içerisinde birbirinden ayrı olan birden fazla kesime ayrılabilir. Bu sebeple bir bölüt birden fazla alt bölütler haline dönüşür ve her bir alt bölüt ayrı olarak değerlendirmeye



Şekil 4. Bölüt ve kontur yönelimlerine ait histogramların oluşturulması ve birleştirilmesi. Üst akış bölütlere, alt akış ise konturlara ait histogramları ifade eder. Soldan sağa doğru, gridin her bir parçasının bölüt ve kontur histogramlarının oluşturulması, histogramların bölütler ve konturlar için birleştirilerek 108 boyuta sahip 2 histogramın oluşturulması ve son olarak bu ikisinin de birleştirilerek 216 boyutlu nihai histogramın elde edilmesi gösterilmiştir.

almır. Benzer şekilde konturlar da alt konturlara ayrılırlar. Gürültüyü önlemek adına belirli bir eşik değerinden küçük alana sahip olan bölütler dikkate alınmaz. Ancak bu işlem konturlar üzerinde uygulanmaz. Şekil 3 bu işlemleri görsel olarak ifade etmektedir.

Bu işlemler sonucunda her bir grid parçası içerisinde bulunan alt bölütler ve alt konturlar için ayrı ayrı yönelim değerleri açı cinsinden hesaplanır. Hesaplanan açı değerleri ile  $15^\circ$ 'lik çözünürlükte olan  $180^\circ$ 'lik bir düzleme oturtulmuş olan histogram oluşturulur. Bu işlem hem konturlar hem de bölütler üzerinde yapıldığından 2 adet histogram elde edilir. Her bir histogramın aralık sayısı  $9 \times (180/15)$ 'dir. Son olarak ise bu iki histogram da birleştirilir ve 216 aralıktan oluşan tek bir histogram elde edilir. Bu işlemler için Şekil 4 incelenebilir.

Son olarak, daha önce oluşturulmuş olan histogramlar sınıflandırma işlemi için nitelik olarak kullanılır.

#### IV. DENEYLER VE DEĞERLENDİRME

Değerlendirme aşamasında İkizler ve diğerleri tarafından oluşturulmuş olan veri kümesi kullanılmıştır [1]. Veri kümesi 6 adet hareket sınıfı içermektedir. Bu sınıflar yakalama, çökme, tekmeleme, koşma, atma ve yürümedir. Veri kümesinde toplam 467 adet resim bulunmaktadır. Resimler rasgele bakış açılarıyla çekilmiş, dağınık arkaplanlara sahip ve farklı çözünürlüklerde. Veri kümesine ait örnek resimler için Şekil 5 incelenebilir. Bu etkenler tanıma işlemini zorlaştırmaktadır.

Hareketlerin sınıflandırılması için Destek Vektör Makineleri (DVM) yöntemi kullanılmıştır. Sınıflandırıcılar oluşturulurken, RBF çekirdeğine sahip DVMler kullanılmış, deneyler tekli çıkarımlı Çapraz Geçerleme (leave-one-out cross validation) kullanılarak değerlendirilmiştir.

Çalışmada 3 farklı sonuç sınıfları arası karışıklık matrisi (confusion matrix) olarak verilmiştir:

- Sadece bölütlere ait 108 boyutlu histogramın kullanılması (Tablo I).
- Sadece konturlara ait 108 boyutlu histogramın kullanılması (Tablo II).
- Bölütlere ve konturlara ait birleşik olan 216 boyutlu histogramın kullanılmasıyla elde edilen sonuçlar (Tablo III).



Şekil 5. Veri kümesinden örnek resimler. Yukarıdan aşağıya sırasıyla koşma, tekmeleme, atma, çökme, yakalama ve yürüme.

Tablo IV'ten görüldüğü üzere sadece kontur bilgisinin kullanılması ile elde edilen sonuçlar nispeten düşüktür (%36.19). Bölüt dağılımları üzerinde yapılan sınıflandırma daha başarılı sonuçlar vermiştir (%45.82). En yüksek başarıyı, önerilen kontur ve bölüt özniteliklerinin birarada kullanılması ile elde edilmiştir (%48.18). Bu noktada, kontur ve bölüt bilgisinin birbirini tamamlayıcı özelliği olduğu görülmektedir. Öte yandan, elde edilen sonuçlar, ara poz çıkarsama yöntemi kullanan diğer yaklaşımlara kıyasla (Aynı veri kümesini kullanan İkizler ve diğerleri tarafından önerilen yöntemin [1] başarıyı %85.1'dir.) yöntemin istenilen düzeyde başarıya sahip olmadığını göstermektedir. Bu durumun en önemli sebeplerinden birisi arkaplan çıkarımı yapılmadığı için insan pozundan bağımsız olan bölüt ve konturların da hesaba katılmasıdır. Özellikle dağınıklığın çok fazla olduğu durumlarda bu etken daha da büyüktür.

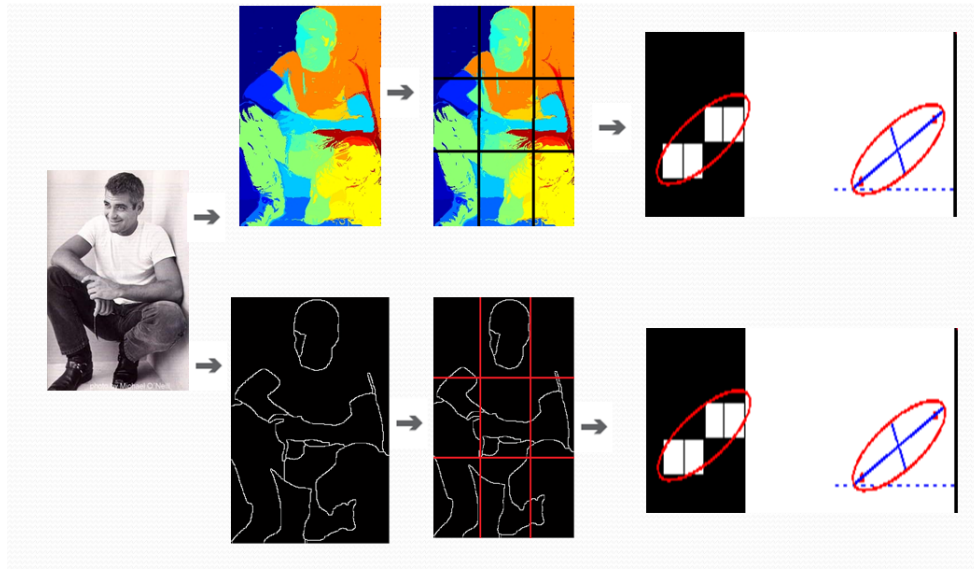
Tablo I. SADECE KONTUR HISTOGRAMLARI KULLANILARAK ELDE EDİLEN KARIŞIKLIK MATRİSİ. SATIRLAR TAHMİN EDİLEN SINIFI, SUTUNLAR ASIL SINIFI GÖSTERMEKTEDİR.

	koşma	atma	yürüme	yakalama	çökme	tekme
koşma	45.5%	11.9%	22.8%	0.0%	9.9%	9.9%
atma	34.7%	28.0%	9.3%	0.0%	12.0%	16.0%
yürüme	25.8%	6.5%	49.5%	1.1%	12.9%	4.3%
yakalama	33.3%	24.1%	18.5%	1.9%	11.1%	11.1%
çökme	14.7%	10.7%	16.0%	1.3%	49.3%	8.0%
tekmeleme	30.4%	20.3%	8.7%	4.3%	10.1%	26.1%

#### V. SONUÇ

Bu çalışmada, resimlerden insan hareketinin tanınması için yeni bir yaklaşım ele alınmıştır. Yaklaşımda, insan pozunun temsil edilmesi için resimlerdeki kontur ve bölütlerin yönelimlerinin histogramları öznitelik olarak kullanılmıştır. Sınıflandırma yöntemi olarak Destek Vektör Makineleri kullanılmıştır. Elde edilen sonuçlar, kontur ve bölüt bilgisinin birlikte kullanımının başarı oranını arttırdığını göstermekte-





Şekil 3. Ön işleminin başlangıç adımı. Üst akış bölütlerin, alt akış ise konturların elde edilmesi ve işlenmesini göstermektedir. Soldan sağa doğru orjinal resim, bölütlerin/konturların elde edilmesi, resmin 9 parçalık grid haline getirilmesi ve yönelimlerin hesaplanması ifade etmektedir.

Tablo II. SADECE BÖLÜT HISTOGRAMLARI KULLANILARAK ELDE EDİLEN KARIŞIKLIK MATRİSİ.

	koşma	atma	yürüme	yakalama	çömelleme	tekme
koşma	46.5%	7.9%	25.7%	4.0%	8.9%	6.9%
atma	34.7%	30.7%	12.0%	2.7%	5.3%	14.7%
yürüme	19.4%	3.2%	69.9%	1.1%	4.3%	2.2%
yakalama	27.8%	18.5%	35.2%	3.7%	1.9%	13.0%
çökme	24.0%	5.3%	4.0%	1.3%	57.3%	8.0%
tekmeleme	18.8%	15.9%	2.9%	1.4%	11.6%	49.3%

Tablo III. KONTUR VE BÖLÜT HISTOGRAMLARININ BİRLİKTE KULLANILMASI İLE ELDE EDİLEN KARIŞIKLIK MATRİSİ.

	koşma	atma	yürüme	yakalama	çömelleme	tekme
koşma	44.6%	5.0%	24.8%	0.0%	10.9%	14.9%
atma	17.3%	34.7%	12.0%	2.7%	4.0%	29.3%
yürüme	18.3%	4.3%	65.6%	4.3%	5.4%	2.2%
yakalama	18.5%	14.8%	33.3%	3.7%	3.7%	25.9%
çökme	16.0%	0.0%	1.3%	1.3%	68.0%	13.3%
tekmeleme	13.0%	7.2%	2.9%	10.1%	8.7%	58.0%

dir. Arkaplan ayrıştırılması kullanılmadan, sadece kontur ve bölüt dağılım bilgisi üzerinden elde edilen bu sonuçlar, umut vericidir. Fakat, resimlerde arkaplanın ayrıştırılmaması sebebiyle başarı oranı yüksek seviyelere ulaşamamıştır. Özellikle dağınıklığın yüksek olduğu resimlerde başarımın düşük olduğu gözlenmiştir. İleriki çalışmalarda arkaplan ayrıştırılması için farklı yöntemler kullanılarak uygulanarak başarımın artırılması hedeflenmektedir.

Tablo IV. YÖNTEMLERİN GENEL BAŞARIM ORANLARI.

Yöntem	Başarım
Kontur	36.19%
Bölüt	45.82%
Kontur ve Bölüt	48.18%

## VI. TEŞEKKÜR

Bu çalışma, 112E149 no'lu TÜBİTAK kariyer projesi tarafından desteklenmiştir.

## KAYNAKÇA

- [1] N. İkozler, R.G. Cinbis, S. Pehlivan and P. Duygulu, *Recognizing actions from still images*, 19th International Conference on Pattern Recognition, ICPR, 2008.
- [2] C. Thureau and V. Hlavac, *Pose primitive based human action recognition in videos or still images*, IEEE Computer Vision and Pattern Recognition, CVPR, 2008.
- [3] W. Yang, Y. Wang, and G. Mori, *Recognizing Human Actions from Still Images with Latent Poses*, IEEE Computer Vision and Pattern Recognition, CVPR, 2010.
- [4] M. Eicher, M.M. Jimenez, A. Zisserman and V. Ferrari, *Articulated Human Pose Estimation and Search in (Almost) Unconstrained Still Images*, ETH Technical Report, September 2010
- [5] A. Agarwal and B. Triggs, *A local basis representation for estimating human pose from cluttered images*, Conference on Computer Vision - January 2006
- [6] P.F. Felzenszwalb and D.P. Huttenlocher, *Graph-based Image Segmentation*, ALADDIN Workshop on Graph Partitioning in Vision and Machine Learning, CMU, 2003.
- [7] J. Shi and J. Malik, *Normalized Cuts and Image Segmentation*, IEEE Transactions on Pattern Analysis and Machine Intelligence, August 2000.
- [8] M. Kass, A. Witkin, D. Terzopoulos, *Snakes: Active contour models*, International Journal of Computer Vision, 1987
- [9] P. Arbelaez, M. Maire, C. Fowlkes and J. Malik, *Contour Detection and Hierarchical Image Segmentation*, IEEE Trans Pattern Anal. Mach. Intell. 33(5):898-916, 2011
- [10] R. Poppe, *A survey on vision-based human action recognition*, Image and Vision Computing, 28 (6). pp. 976-990, 2010.