Yüksek Boyutlu Öznitelik Uzayında Hareket Tanıma

Action Recognition in a High-Dimensional Feature Space

Hande Adiguzel, Hayrettin Erdem, Hakan Ferhatosmanoğlu, Pinar Duygulu
Bilgisayar Mühendisliği Bölümü
Bilkent Üniversitesi
Ankara, Türkiye
{adiguzel, hayrettin, hakan, duygu}@cs.bilkent.edu.tr

Özetçe— Hareket analizi ve tanıma bilgisayarla görü alanında önemli ve zorlayıcı bir alan olarak görülmektedir ve insan hareketlerini tanımlamak için değişik teknikler kullanılmaktadır. Bu makalede zaman-mekansal özniteliklerin kullanılması tercih edilmiştir. Aynı zamanda, zaman-mekansal özniteliklerin oldukça yüksek ölçülü olmasından kaynaklanan boyutluluk sorunu ortaya çıkmaktadır ve geleneksel benzerlik ve eşleme algoritmaları bu problemin üstesinden gelememektedir. Bu sebeple, makalede yüksek boyutlu özniteliklerin eşleştirilmesi için farklı bir yöntem uygulanmıştır ve imgelerin eşleştirilmesi yöntemiyle videolar değişik hareketlere sınıflandırılmıştır.

Anahtar Kelimeler — Hareket Tanıma, İnsan Hareketlerinin Sınıflandırılması, Boyutluluk Sorunu, Yüksek Ölçülü Uzay

Abstract— Analyzing and interpreting human actions is an important and challenging area of computer vision. Different solutions are used for representing human actions; we prefer to use spatio-temporal interest points for motion descriptors. Besides, the space-time interest point feature space is considerably high-dimensional and it is hard to eliminate the curse of dimensionality with traditional similarity functions. We apply a matching based approach for high dimensional feature space that matches sequences to classify actions.

Keywords — Action Recognition, Recognizing Human Motion, Curse of Dimensionality, High-Dimensional Space

I. GİRİŞ

İnsan Hareketlerinin incelenmesi ve sınıflandırılması bilgisayarla görü alanında oldukça önemli ve ilgi çekici bir çalışma konusudur. İlgi çekici bir alan olmasının en önemli sebebi birçok değişik uygulamada kullanılmasıdır. Bunlardan bazıları tarama, gözetleme, video dinleme, atletik performans analizi, içerik temelli imge biriktirme ve geri getirme gibi sıralanabilir [1].

Hareket analizi için kullanılan geleneksel yöntemlerden bazıları optik akışı [2] ve öznitelik takibidir [3]. Fakat, bu yöntemler bazı durumlarda yetersiz kalmaktadır. Optik akış yöntemi harekette ani değişimler olduğunda başarısız olmaktadır. Öznitelik takibi ise imgelerin zamana bağlı değişimini kullandığı için görünümdeki değişimler örneğin 2

farklı cismin kesişmesi veya ayrılması gibi durumlarda yetersiz kalmaktadır [1].

Bu ve bunun gibi problemleri farklı yaklaşımlar ile çözmek mümkündür. Eylemleri hem zaman hem de mekansal tanım kümesinde temsil ettiğimizde eksiksiz bir tanım elde etmiş oluruz. Mekansal uzayda belirgin yerel çeşitlilik gösteren noktalar içerdikleri temsilci bilgiler sayesinde imgeleri tanımlamak için kullanılmaktadır. Bu noktalara özel tanımlayıcı noktalar denmektedir [4]. Başka bir makale olan [5]'te ise Laptev tanımlayıcı noktalar kavramını geliştirip zaman-mekansal özniteliğini elde etmiştir ve bunları videolardaki ilginç olayları ve eylemleri tespit etmek için kullanmıştır. Bu zaman-mekansal noktalar (STIP) eylemlerin tespitini öznitelik takibi, imge bölütleme veya optik akışı hesabı olmadan yapmaktadır. Bizim çalışmamızda da zaman-mekansal öznitelik vektörleri kullanılmıştır.

Hareketleri sınıflandırmada videonun imgelerine ait zaman-mekansal öznitelikler eşleştirilmiştir. Eşleştirme algoritması olarak [6] makalesinden farklı bir yöntem kullanılmıştır. Bunun sebebi ise zaman-mekansal özniteliklerin boyutunun yüksek (162) olmasıdır. Geleneksel eşleştirme ve benzerlik algoritmaları yüksek boyutlu uzaylarda işlevsiz kalmaktadır [7]. Bu fenomene boyutluluk problemi denmektedir, aynı zamanda sınıflandırma ve öbekleme konularında problemlere yol açmaktadır. Bu nedenle, çalışmamızdaki amaç insan hareketlerini doğru bir biçimde temsil ederken aynı zamanda yüksek boyutlu uzaylar için uygun eşleştirme ve benzerlik algoritmaları kullanmak ve boyutluluk sorununu ortadan kaldırmaktır.

Sıradaki bölüm ilgili çalışmalardan bahsederken sonraki bölüm veri kümesini ve çıkarılan öznitelikleri açıklamaktadır. Ardından, tanımlayıcı noktaları eşleştirmek için önerdiğimiz yöntem anlatılmaktadır. Son olarak ise, deneyler ve sonuçlar bölümleri mevcuttur.

II. GEÇMİŞ ÇALIŞMALAR

Zaman-mekansal öznitelikler başka hareket tanıma çalışmalarda da kullanılmıştır. Bazı çalışmalar zaman-mekansal öznitelikleri BoW(Bag-of-Words) tekniği ile kullanmaktadır [8, 9, 10]. Fakat bu teknik, hareketin yönünü, zamansal bağımlılık ve ilişkileri saklamamaktadır ve bu

özellikler hareket tanıma için oldukça önem arz etmektedir [11, 12].

Diğer bir yöntem olan optik akış videodaki cisimlerin hareket yönünü temsil etmektedir ve birçok çalışma bu yöntemi hareket tanımda kullanmaktadır [13, 14]. Fakat, optik akış videodaki görüntü değişimlerine karşı aşırı duyarlıdır ve ani değişimlerde başarısız olmaktadır [15].

Hareket tanımda başka bir yöntem ise insan bedeninin pozlarının tespiti. Bu yöntemin arkasındaki mantık, bir hareketin tüm videoyu incelemiden tek bir imgeden de anlaşılabilirliği yönündeki sezgidir [16]. Poz bilgisini kullanan çalışmaların çoğu [17, 18, 19] insan vücudunun bileşenlerinin sıklık grafiğini çıkarmaktadır. Fakat bu yöntem bileşenler arasındaki geometrik ilişkiyi göz ardı etmektedir.

Geçmiş çalışmalarda hareketi temsil etmede kullanılan BoW veya bileşenlerin sıklık grafiği yöntemleri yukarıda açıkladığımız gibi birtakım problemlere yok açmaktadır. Bizim çalışmamızdaki amaç özneteliklerin sınıflandırmada boyutsal indirgemeye gerek olmadan doğrudan kullanılmasıdır. Bu sebeple de yüksek boyutlu uzaylara uygun bir eşleme tekniğinin uyarlanması gerekmektedir.

III. VERİ KÜMESİ VE ÖZNİTELİK UZAYI

Veri kümesi olarak KTH [20] kullanılmıştır. Bu veri kümesi 6 farklı insan hareketi (yürüme, tempolu koşma, koşma, boks, el sallama ve el çırpma) içermektedir (Şekil 1). Hareketler 25 değişik insan tarafından 4 farklı senaryoda gerçekleştirilmiştir. Senaryolar dışarıda, değişik ölçeklerde dışarıda, farklı kıyafetler ile dışarıda ve içeride koşullarını içermektedir. Toplamda 2391 farklı video bulunmaktadır.

Şekil 1. KTH veri kümesinden değişik kare örnekleri

Öznetelik olarak önceden de bahsettiğimiz gibi zaman-mekansal (STIP) tanımlayıcı noktalar [6] kullanılmıştır. Şekil 2'de bir yürüme videosundan hesaplanan öznetelikler görülmektedir. Öznetelikler çıkarılırken çoklu zaman-mekansal ölçüler kullanılmaktadır. Sonuç olarak 2 farklı tanımlayıcı öznetelik hesaplanmaktadır. Bunlardan ilki yönlü gradyanların sıklık grafiği (HOG) ikincisi ise optik akışın sıklık grafiğidir (HOF).

Şekil 2. Örnek bir yürüme hareketi videosundan hesaplanan zaman-mekansal öznetelikler.

IV. ÖNERİLEN YÖNTEM

Önerilen yöntemde öncelikli olarak hesaplanan özneteliklerdeki tekrarlı motiflerin bulunması için k-ortalamlar yöntemi uygulanmıştır. Ardından, kalabalık öbeklerin merkezleri hesaplanarak hareket sınıfları bu merkezlerle temsil edilmiştir. Her imge grubu için hesaplanan öbek merkezlerine k-n-eşleştirme [21] yöntemi uygulanmıştır ve eşlenen ikililer arasındaki uzaklıklar hesaplanmıştır. Bir test videosu ile eğitim kümesindeki videolar arasındaki bütün uzaklıklar hesaplandıktan sonra test videosu k-en-yakın eğitim videosunun oy çoğunluğuna göre sınıflandırılmıştır.

A. Eşleme Yöntemi

Test kümesindeki videoların sınıflandırılması için eğitim kümesinden elde edilen öznetelikleri test kümesindekiler ile eşleştirebiliriz. Fakat yaklaşık olarak her imgeden 500 tane 162 boyutlu STIP elde edilmektedir ve bu durumda eşleştirme oldukça zaman alıcı bir işlem haline gelmektedir. Aynı zamanda arka plandan çıkarılanlar gibi elde edilen özneteliklerin hepsi tanımlayıcı özellikte olmayabilir. Laptev [5] makalesinde tanımlayıcı örüntülerin tespit edilmesi için bir yöntem önermektedir.

Bir hareketin tekrarlanan eylemlerinin bulunması için özneteliklere k-ortalamlar yöntemi uygulanabilir. Çıkan sonuçlarda kalabalık öbeklerin o harekete ait tekrarlanan eylemlerinin özneteliklerini içermesi beklenmektedir. Şekil 3'ten de görüleceği gibi yürüme hareketinin en kalabalık 4 öbeğinin sonuçları en tanımlayıcı öznetelikleri içermektedir.

Bizim çalışmamızda k 15 olarak ayarlanmıştır ve en kalabalık 5 öbek merkezi incelenmiştir. Bu yöntem tanımlayıcı olmayan özneteliklerin de elenmesini sağlamaktadır. Test ve eğitim kümesindeki her video için en kalabalık 5 öbek merkezi hesaplandıktan sonra merkezler eşleştirilerek videolar arasındaki uzaklıklar hesaplanmıştır.

Eşleştirme algoritması olarak yüksek boyutlu uzaylara uygun k-n-eşleştirme algoritması kullanılmıştır. Şekil 3'ün ikinci satırında da görüldüğü gibi test videosundaki öbek merkezleri eğitim videosundaki benzer eylemleri temsil eden öbek merkezlerine eşleştirilmiştir. Son olarak elde edilen uzaklıklardan her test videosunun k-en-yakın eğitim videosu bulunmuştur ve test videosu sonuçlardaki oy çoğunluğuna göre bir sınıfa atanmıştır.

Şekil 3. İlk satır: Eğitim kümesindeki bir videonun öbekleme sonuçları. İşaretlenen noktalar en kalabalık öbeklere denk gelmektedir. İkinci satır: Örnek bir test videosunun ilk satırdaki eğitim videosuyla eşleştirme sonuçları.

B. K-n-eşleme

Birçok çalışma, benzerlik bulma görevini, yakın komşu bulma problemi olarak ele almaktadır. Yakın komşu modeli, sabit bir özellikler kümesi üzerinde, sorgu ve diğer veri

nesneleri arasındaki uzaklığın bulunmasına dayanmaktadır. Fakat benzerlik arama görevinin bu tür bir yaklaşımla yerine getirilmesi birtakım dezavantajları beraberinde getirmektedir. Bunlardan ilki, sorgu ve veri nesneleri arasındaki kısmi benzerliklerin gözden kaçırılmasıdır. Örneğin, birtakım resim nesneleri arasında, renk, desen, biçim gibi alt özellik kümelerine dikkat edilmeksizin, tüm özelliklerin oluşturduğu küme üzerinde uzaklık bulunması, yanlış sonuçlara neden olmaktadır. Bir diğer dezavantaj ise uzaklığın genellikle birkaç boyutta görülen yüksek benzersizlikten dolayı etkilenmesidir. Yani, iki nesne, boyutların birçoğunda birbirine benzerken, birkaç boyutta görülen yüksek benzersizlikler, bu iki nesnenin birbirine benzemediği sonucuna varılmasına neden olmaktadır. Aksiyon tanımlama görevinde de, yukarıda belirtilen durumlar gözlenmektedir. Bu eksiklikleri giderebilmek için, [21] k-n-eşleme yöntemini önermektedir.

K-n-eşleme yöntemi, sorgu nesnesini veri nesneleriyle en iyi eşlendikleri n boyutta seçer (n sayısı en iyi eşlemeyi sağlayacak şekilde dinamik olarak belirlenir). Sonrasında, sorgu nesnesine en çok benzeyen k tane veri nesnesi seçilir. Algoritmanın işleyişi temel olarak şu şekildedir: Öncelikle, sorgu ve her bir veri nesnesi arasındaki n-eşleme farkı bulunur. Verilen bir çift nesne için bu fark, basit olarak her boyuttaki değerlerin mutlak farkının oluşturduğu bir vektör içerisindeki n. küçük farktır. Sonraki adımda, sorgu nesnesiyle en çok benzeyen veri nesneleri seçilir. Fakat sorgu nesnesiyle birebir uyuşan nesneler bulunamayabileceği için bir eşik değeri belirlenir. Bu eşik değeri k.en küçük n-eşleme değeri olarak seçilir ve bu da sorgu nesnesine en çok benzeyen k tane veri nesnesinin bulunması anlamına gelmektedir. Fakat böyle bir algoritmanın kullanılması, yüksek bir hesaplama maliyeti gerektirmektedir. Çünkü n-eşleme farkı hesaplanırken, her bir veri nesnesinin her boyutuna erişim gerekmektedir. Bu yüksek hesaplama maliyetinden kurtulmak için, [21] kullanılan veri nesnelerini yeniden düzenleyip, ardından k-n-eşleme algoritmasını uygulamaktadır. Yani, bu düzenleme, tüm nesnelerin değerlerinin oluşturduğu yapıyı, her boyutta bulunan değerlerin küçükten büyüğe sıralı olduğu bir yapıya dönüştürmektedir. Sonunda, boyut sayısı kadar sıralı listeler elde edilmiş olur. Şekil 4 örnek bir veri setini ve verilerin yeniden düzenlenmiş halini göstermektedir.

ID	d ₁	d ₂	d ₃
1	0.2	2.5	1.0
2	7.2	1.0	3.1
3	1.4	1.8	4.7
4	1.0	5.2	2.2
5	3.0	6.0	1.6

d ₁	d ₂	d ₃
1, 0.2	2, 1.0	1, 1.0
4, 1.0	3, 1.8	5, 1.6
3, 1.4	1, 2.5	4, 2.2
5, 3.0	4, 5.2	2, 3.1
2, 7.2	5, 6.0	3, 4.7

Şekil 4. Örnek veri kümesi ve yeniden düzenlemesi

Yeni düzenlenen veri kümesinde algoritma şu şekilde çalışmaktadır: Öncelikle, sorgu nesnesini ifade eden her boyuta ait değerler, yeni düzenlenen veri kümesi üzerinde bulunur. Sonrasında, her boyut için, o boyuta ilişkin sorgu değerine en yakın değere sahip olan nesneler, bu değerlerin mutlak farkıyla birlikte bir dizide tutulur. Bu dizide, bir nesne,

ilk kez n defa görüldüğünde, o nesne ilk n-eşleme nesnesi olur. Benzer şekilde, n defa görülme sıklığına ulaşan ilk k tane nesne, sorgu nesnesine en benzer k tane nesne olarak algoritma tarafından döndürülür.

K-n-eşleme yöntemi sayesinde, uygun bir n değeri için, sorgu nesnesine benzer nesneler bulunmaktadır. Fakat uygun n değerinin bulunması, her zaman açık olmayabilir. Bu yüzden, [21] uygun n değerini saptamak için, frequent k-n-match yöntemini önermektedir. Bu yöntem, uygun n değerini bulurken, belirli bir aralıktaki değerleri deneyip, her değer için elde edilen sonuç kümesi üzerinde birtakım hesaplamalar yapmaktadır. Başka bir deyişle, ilk adımda, seçilen aralıktaki her n değeri için elde edilen sonuç kümesi kaydedilir. Sonrasında, tüm sonuç kümeleri içerisinde, gözükme sıklığı, seçilen değer aralığında olan ilk k tane nesne, sorgu nesnesine en benzer sonuçlar olarak bulunur. Aralık seçimi, kullanıcılar tarafından yapılabileceği gibi, sistem tarafından otomatik olarak belirli bir aralıktaki seçilebilmektedir. Değer aralığı seçilmesinde, dikkat edilmesi gereken birtakım noktalar vardır. Bunlardan ilki, aralığın alt sınırını temsil eden değerin çok küçük seçilmemesidir. Çünkü seçilen bu değer, bir nesneyi kısmi olarak iyi ifade etmeyebilir ya da seçilen boyutlar, verideki gürültüden dolayı seçilmiş olabilir. Yine benzer şekilde, aralığın üst sınırını temsil eden değer, çok büyük olmamalıdır. Çünkü boyut sayısı arttıkça, kısmi benzerlikten faydalanma azalır ve benzerlik araması, büyük benzersizlik farklarının görüldüğü boyutlardan etkilenebilir.

V. DENEYLER

Veri kümesi 2 parçaya ayrılmıştır, ilki eğitim kümesidir ve 12 farklı kişi tarafından 4 farklı koşulda gerçekleştirilen toplam 288 hareket videosu içermektedir. İkincisi ise 13 farklı kişi tarafından 4 farklı koşulda gerçekleştirilen toplam 312 video içermektedir. Eğitim ve test kümesindeki kullanılan koşullar aynıdır fakat kişiler tamamen farklıdır.

Temel sistem geleneksel bir yöntem olan k-en-yakın eşleştirmeyi kullanmaktadır. Tablo 1'de hata matrisinden de görüleceği üzere toplam tanıma oranı %55.8'dir. El çırpma ve sallama sınıflarının tanıma oranlarının yüksek olmasının sebebi diğer sınıflardaki videolardan tespit edilen öbek merkezlerinin aykırı değerlere sahip olması ve bunların hesaplanan uzaklıklara hakim olması sonucunda videoların yanlış sınıflandırılmasıdır. Bu problem [21] makalesinde de açıklanmaktadır. Gerçekte birbirine yakın olan vektörler birkaç aykırı boyut değeri yüzünden birbirine uzak olarak hesaplanmaktadır.

Tablo 2 makalede önerilen yöntem ile bulunan tanıma sonuçlarının hata matrisini göstermektedir. Toplam tanıma oranı %75'dir. Başarı oranı çok yüksek olmasa da temel sistemin başarısını geçmektedir ve ileriki çalışmalar için umut vadetmektedir. Aynı zamanda temel karışıklıklar yürüme, koşu ve tempolu koşu sınıfları arasında yaşanmıştır. Bunun sebebi olarak hareketlerin benzerliğini gösterebiliriz.

Aynı zamanda önerilen sistemi 13 farklı kişiyle fakat tek koşulda test ettiğimizde başarı oranının %80'e yükseldiği gözlemlenmiştir.

	Yürüme	Tempolu koşu	Koşu	Boks	El sallama	El çırpma
Yürüme	0.65	0	0	0	0.20	0.15
Tempolu koşu	0	0.15	0	0	0.10	0.75
Koşu	0.25	0	0.35	0	0.05	0.35
Boks	0	0	0	0.50	0.15	0.35
El sallama	0.05	0	0	0	0.80	0.15
El çırpma	0	0	0	0	0.10	0.90

Tablo. 1. 13 farklı kişinin 4 farklı koşulda gerçekleştirdiği hareketlerin temel sistem kullanılarak tanınması sonucunda hesaplanan hata matrisi.

	Yürüme	Tempolu koşu	Koşu	Boks	El sallama	El çırpma
Yürüme	0	0.13	0.75	0	0	0.12
Tempolu koşu	0	0.25	0.21	0	0	0.54
Koşu	0.02	0.92	0.02	0	0	0.04
Boks	0.79	0.02	0	0.19	0	0
El sallama	0.06	0.08	0	0.13	0.73	0
El çırpma	0.10	0.05	0	0.75	0.10	0

Tablo. 2. 13 farklı kişinin 4 farklı koşulda gerçekleştirdiği hareketlerin önerilen k-n-eşleştirme sistemi kullanılarak tanınması sonucunda hesaplanan hata matrisi. Kullanılan değer kümesi [10, 60]'tır.

VI. ÖZET VE TARTIŞMA

Çalışmamızda zaman-mekansal öznitelikler kullanılarak hareket videoları tanımlanmıştır ve videoların sınıflandırmasında yüksek boyutlu uzaylara uygun k-n-eşleştirme yöntemi tercih edilmiştir. Deneylerden de gözlemleneceği gibi geleneksel yöntemler kullanan temel sistemin başarı oranları oldukça düşüktür ve önerilen yöntem tanıma oranlarını yükseltmiştir. Değişik insanlar tarafından değişik koşullarda gerçekleştirilen hareketlerin k-n-eşleştirme yöntemiyle hesaplanan başarı oranı %75'dir. Tanıma oranı çok yüksek olmamakla beraber gelecek çalışmalar için umut vadedicidir. Başarı oranlarının artırılması için değişik ve birden fazla öznitelik kullanılabilir veya eğitim kümesindeki videoların sayısı artırılabilir.

KAYNAKÇA

- [1] J. K. Aggarwal and Q. Cai. Human motion analysis: a review. *Computer Vision and Image Understanding*, 73(3):428-440, 1999.
- [2] Barron, J., Fleet, D., and Beauchemin, S. 1994. Performance of optical flow techniques. *International Journal of Computer Vision*, 12(1):43-77.
- [3] Smith, S. and Brady, J. 1995. ASSET-2: Real-time motion segmentation and shape tracking. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(8):814-820.
- [4] Förstner, W.A. and Gülch, E. 1987. A fast operator for detection and precise location of distinct points, corners and centers of circular features. In *Proc. Intercommission Workshop*

of the Int. Soc. for Photogrammetry and Remote Sensing, Interlaken, Switzerland.

- [5] "On Space-Time Interest Points" (2005), I. Laptev; *International Journal of Computer Vision*, vol 64, number 2/3, pp.107-123.
- [6] Laptev, I., & Lindeberg, T. (2006). Local descriptors for spatiotemporal recognition. In *Lecture notes in computer science* (Vol. 3667). Spatial coherence for visual motion analysis, first international workshop, SCVMA 2004, Prague, Czech Republic, 15 May 2004. Berlin: Springer.
- [7] Beyer, K.; Goldstein, J.; Ramakrishnan, R.; Shaft, U. (1999). "When is "Nearest Neighbor" Meaningful?". *Proc. 7th International Conference on Database Theory - ICDT'99*. LNCS 1540: 217-235.
- [8] P. Dollar, V. Rabaud, G. Cottrell, and S. Belongie. Behavior recognition via sparse spatio-temporal features. In *VS-PETS*, 2005.
- [9] J. C. Niebles, H. Wang, and L. Fei-Fei. Unsupervised learning of human action categories using spatial-temporal words. *International Journal of Computer Vision*, 79(3):299-318, 2008.
- [10] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld. Learning realistic human actions from movies. In *CVPR*, 2008.
- [11] A. Kovashka and K. Grauman. Learning a hierarchy of discriminative space-time neighborhood features for human action recognition. In *CVPR*, 2010.
- [12] A. Ta, C. Wolf, G. Lavoue, A. Baskurt, and J. Jolion. Pairwise features for human action recognition. In *ICPR*, 2010.
- [13] Y. Wang, P. Sabzmeydani, and G. Mori. Semi-latent dirichlet allocation: a hierarchical model for human action recognition. In *ICCV Workshop on Human Motion*, 2007.
- [14] A. Fathi and G. Mori. Action recognition by learning mid-level motion features. In *CVPR*, 2008.
- [15] P. Turaga, R. Chellappa, V. S. Subrahmanian, and O. Udrea. Machine recognition of human activities: a survey. *Circuits and Systems for Video Technology*, 18(11):1473-1488, 2008.
- [16] S. Baysal, M. C. Kurt, and P. Duygulu. Recognizing human actions using key poses. In *ICPR*, 2010.
- [17] N. Ikizler and P. Duygulu. Human action recognition using distribution of oriented rectangular patches. *Image and Vision Computing*, 27(10), 2009.
- [18] K. Hatun and P. Duygulu. Pose sentences: a new representation for action recognition using sequence of pose words. In *ICPR*, 2008.
- [19] C. Thureau and V. Hlavac. Pose primitive based human action recognition in videos or still images. In *CVPR*, 2008.
- [20] "Recognizing Human Actions: A Local SVM Approach", C. Schuld, I. Laptev, B. Caputo; In *Proc. ICPR 2004*, Cambridge, UK.
- [21] Anthony K. H. Tung, Rui Zhang, Nick Koudas, Beng Chin Ooi. Similarity search: a matching based approach, *Proceedings of the 32nd international conference on Very large data bases*, September 12-15, 2006, Seoul, Korea