

Çoklu Sorgu Yöntemiyle Mobil Görüntü Arama Mobile Image Search Using Multi-Query Images

Fatih Çalışır¹, Muhammet Baştan², Uğur Güdükbay¹, Özgür Ulusoy¹
¹Bilgisayar Mühendisliği Bölümü, Bilkent Üniversitesi, Ankara, Türkiye
²Bilgisayar Mühendisliği Bölümü, Turgut Özal Üniversitesi, Ankara, Türkiye

fatih.calisir@bilkent.edu.tr, mbastan@turgutozal.edu.tr, gudukbay@cs.bilkent.edu.tr, oulusoy@cs.bilkent.edu.tr

Özetçe—Mobil cihaz teknolojisindeki son gelişmeler, cep telefonlarını yüksek çözünürlüklü kameraları olan ve hızlı görüntü işleyen güçlü cihazlar haline getirmiştir. Normal bilgisayarlara göre kullanıcı ile çok daha fazla ve kolay etkileşim imkanı sağlayan cep telefonları, bu özellikleri sayesinde, daha zengin görüntü tabanlı arama yapmayı da sağlayabilir. Bir nesnenin tek bir görüntüsü ile arama yapmak yerine, cep telefonu ile gerçek zamanlı olarak, farklı uzaklık ve açılardan çekilen görüntüleri ile nesne hakkında daha fazla görsel bilgi edinilebilir ve böylece sorgu sonuçları iyileştirilebilir. Bu çalışmada, aranacak nesnenin mobil cihaz kamerasıyla elde edilen birden çok görüntüsü kullanılarak görsel nesne sorgu sonuçlarının iyileştirilmesi amaçlanmaktadır. Bunun için, görsel kelimeler (visual words) ile birden çok sorgu resminden edinilen bilgiler; sorgu öncesi görsel kelime histogramı oluştururken (early fusion) veya tek tek her resim ile elde edilen sorgu sonuçları birleştirilmiştir (late fusion). Nesne görüntülerinden oluşan bir veri seti üzerinde yapılan kapsamlı deneylerle, çoklu sorguların, tek resimle yapılan sorgulardan daha iyi sonuç verdiği gösterilmiştir.

Anahtar Kelimeler — mobil görüntü araması; içerik temelli resim araması; çoklu sorgu; sorgu birleştirme, görsel kelime histogramı.

Abstract—Recent advances in mobile device technology have turned the mobile phones into powerful devices with high resolution cameras and fast processing capabilities. Having more user interaction potential compared to regular PCs, mobile devices with cameras can enable richer content-based object image queries: the user can capture multiple images of the query object from different viewing angles and at different scales, thereby providing much more information about the object to improve the retrieval accuracy. The goal of this paper is to improve the mobile image retrieval performance using multiple query images. To this end, we use the well-known bag-of-visual-words approach to represent the images, and employ early and late fusion strategies to utilize the information in multiple query images. With extensive experiments on an object image dataset with a single object per image, we show that multi-image queries result in higher average precision performance than single image queries.

Keywords — Mobile visual search; content-based image search; multi-query; combining queries; bag of visual words.

I. GİRİŞ

Akıllı cep telefonlarının teknolojik açıdan gelişmesi ve insan hayatında vazgeçilmez bir yer edinmesi sebebiyle içerik tabanlı resim arama sistemleri mobil ortamda da kullanılmaya başlanmıştır. Kullanıcıların cep telefonlarıyla çektikleri bir resim hakkında bilgi edinmek istemeleri, benzer resimleri görmek istemeleri gibi kullanım amaçları mobil ortamda resim arama sistemlerinin önemini artırmaktadır.

İçerik tabanlı resim arama sistemlerinde kullanılan en yaygın yöntemlerden birisi görsel kelimelerin kullanılmasıdır. Bu yöntem resimlerdeki ilgi noktalarının bulunmasına, bu noktaların çevresinden çıkarılan özniteliklerin nicelenmesinden oluşan bir kod tablosu oluşturulmasına ve bu görsel kelimeler kullanılarak farklı iki resim arasında bir benzerlik değeri bulunmasına dayanmaktadır. Bu çalışmada da bu yöntem kullanılmıştır.

II. MOBİL ORTAMDA RESİM ARAMA

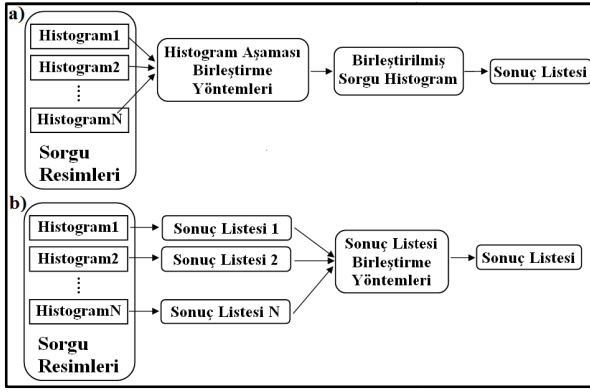
Mobil ortamın kendine has zorluklarının olması mobil ortamda resim araması alanında araştırmaların yoğunlaşmasına neden olmuştur. Tüm resim arama sistemlerinde kullanıcıya verilen sonuç listesindeki resimlerin sorgu resmine benzerliği, sistemin başarısı ve kullanıcı tatmini açısından önemli bir rol oynamaktadır. Buna ek olarak sorgu resmi çekildikten sonuçların kullanıcıya verilmesine kadar geçen sistemin çalışma süresi de kullanıcı tatmini açısından önemlidir. Bu nedenle mobil ortamda resim aramasında bazı çalışmalar süreyi kısa tutmak adına tek sorgu resmi kullanırken, bazı çalışmalar verilen sonuçların sorgu resmine daha benzer olmasını amaçladıklarından dolayı çoklu sorgu resmi kullanmaktadır.

A. Tek Sorgu Resmi Kullanan Sistemler

Tek sorgu resmi kullanan sistemler çalışma süresinden kazanırken, sorgu resminden kısıtlı bilgi almaktadır. Bu sistemler çekilen resimlere benzer resimler listelemek ya da çekilen resim hakkında bilgi vermek gibi amaçlar için kullanılmaktadır.

B. Çoklu Sorgu Resmi Kullanan Sistemler

Çoklu sorgu resmi kullanan sistemler, sorgulanacak nesnenin farklı açılardan resimlerini çekerek daha fazla bilgi sahibi olabilmeleri nedeniyle sonuç listesindeki resimlerin sorgulanan nesneye benzerliğinin daha fazla olması beklenmektedir. Çoklu sorgu resmi kullanımının en önemli noktalarından biri elde edilen sonuçları ya da sorgu resimlerini birleştirme yöntemidir [1]. Bu amaç için kullanılan ilk yaklaşımda, sorgu resimlerinin görsel kelimelere göre çıkartılan histogramları her bir sütun için ortalama değer alma, en büyük ya da en küçük değeri seçme gibi yöntemlerle birleştirilip tekrar tek bir sorgu histogramı oluşturulur (Şekil 1 a) [2]. İkinci yaklaşımda ise, her bir sorgu resmi ayrı ayrı sorgulandıktan sonra elde edilen sonuç listeleri, listedeki resimlerin sorgu resmiyle benzerlikleri de kullanılarak birleştirilir (Şekil 1 b) [3].



Şekil 1. Çoklu sorgu kullanılan sistemlerde birleştirme yöntemleri

III. ÖNERİLEN SİSTEM

Bu çalışmada önerilen sistem diğer birçok resim arama sisteminde olduğu gibi görsel kelimeler yaklaşımını kullanmaktadır.

A. İlgili Noktaları Bulunması ve Öznitelik Çıkarılması

Literatürde çok sayıda ilgi noktası ve öznitelik yöntemi mevcuttur. Bu çalışmada OpenCV (<http://opencv.org>) kütüphanesindeki yedi farklı ilgi noktası; HARRIS, SURF, FAST, ORB, BRISK, FREAK, DENSE bulma ve beş farklı öznitelik çıkarma; BRIEF, BRISK FREAK, SIFT, SURF yöntemi kullanılmıştır. Bu yöntemlerin tüm birleşim ihtimalleri için deneyler yapılmıştır. İlgili noktası için HARRIS ve SURF yöntemleri, SIFT öznitelikleriyle en iyi sonuçları vermiştir. Görüntülerdeki farklı özellikteki ilgi noktalarının birleştirilerek kullanılması performansı artırdığından, HARRIS ve SURF ilgi noktalarından çıkarılan öznitelikler, histogramları birleştirilerek kullanılmıştır.

Görsel kelimeler çıkarılırken önemli noktalardan birisi de k ortalamalar kümeleme yönteminde seçilecek olan ve her bir kümeye atanacak öznitelik sayısını gösteren k değerinin seçimidir. Bu nedenle 1000'den 10000'e kadar 1000 artış miktarıyla on farklı k değeri ile resim veritabanı üzerinde deneyler yapılmış ve en iyi sonucu veren 3000 değeri seçilmiştir.

B. Benzerlik Ölçütleri

Sorgu resmine benzer resimlerin bulunabilmesi için veritabanındaki resimler ile sorgu resimlerinin benzerlikleri ölçülmelidir. Bu amaçla histogramlar arasındaki benzerlik değerini veren yöntemler kullanılmıştır [2][4]. Benzerlik ölçütleri ve hesaplanışları Tablo 1'de görülmektedir. Sorgu resminin histogramı q , karşılaştırılan veritabanı resminin histogramı ise d ile sembolize edilmiştir. Yapılan deneyler sonucu en iyi sonucu *Küçük Büyük Oranı* yöntemi verdiği için deney sonuçlarında bu benzerlik ölçütü kullanılmıştır.

Benzerlik Ölçütü	Hesaplanışı
Normalize Korelasyon	$\frac{\sum_i q_i d_i}{\sqrt{\sum_i q_i^2} \sqrt{\sum_i d_i^2}}$
Normalize Histogram Kesişimi	$\sum_i \min\left(\frac{q_i}{\sum_i q_i}, \frac{d_i}{\sum_i d_i}\right)$
Histogram Kesişimi	$\frac{\sum_i \min(q_i, d_i)}{\min(\sum_i q_i, \sum_i d_i)}$
Nokta Çarpımı	$\sum_i q_i d_i$
Küçük Büyük Oranı	$\frac{\sum_i \min(q_i, d_i)}{\sum_i \max(q_i, d_i)}$

Tablo 1. Benzerlik ölçütleri ve hesaplanışları

C. Çoklu Sorguların Histogram Aşamasında Birleştirme

Bu yöntemler çoklu sorgu resimlerini, histogramları çıkarıldıktan sonra arama işlemi gerçekleşmeden birleştirerek tüm sorgu resimlerinin bilgilerini barındıran tek bir histogram oluşturur. Bu çalışmada iki histogramı birleştirirken karşılıklı sütunlardan en büyük (*HistogramEnBüyük*), ortalama (*HistogramOrtalama*) ve toplam (*HistogramToplam*) değeri seçmek üzere üç farklı yöntem denenmiştir [2][3]. Bunlara ek olarak, bu yöntemler kullanılan veritabanındaki histogramlara bakılarak verilen önem oranına göre adapte edilerek (*Histogram En Büyük Ağırlıklı*, *Histogram En Küçük Ağırlıklı*, *Histogram Ortalama Ağırlıklı*) sunulmuştur. Bu yöntemler için önem oranı, her bir histogram sütunu için, o sütundaki değer tüm sütunlardaki değerlerin toplamına bölünmesiyle elde edilmiştir.

D. Çoklu Sorguların Sonuç Listelerini Birleştirme

Bu yöntemler çoklu sorgu resimleri arama işlemine girdikten sonra elde edilen sonuç listelerinin birleştirilmesi ile tek bir sonuç listesi vermektedir. Beş farklı yöntem kullanılmıştır [1][2].

1) *Benzerlik Değeri Toplamı*: Bu yöntemde elde edilen sonuç listelerindeki aynı resimlerin benzerlik değerleri toplanarak tekrar sıralama yapılır ve tek liste elde edilir.

2) *Ağırlıklı Birleştirme*: Bu yöntemde her bir resim için, benzerlik değeri bütün listelerdeki benzerlik değerlerinin toplamına bölünerek bir önem oranı bulunur. Daha sonra

resmin benzerlik değeri bu önem oranı ile çarpılır ve sıralama yeniden yapılır.

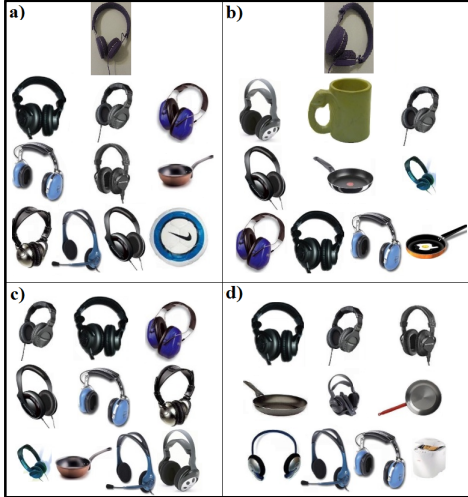
3) *Sıralama Toplamı*: Bu yöntemde elde edilen listelerdeki aynı resimlerin sıralama değerleri toplanarak tekrar sıralama yapılır

4) *En Büyük Benzerlik Değeri*: Elde edilen sonuç listelerindeki aynı resimlerin benzerlik değerlerinden büyük olan alınarak tekrar sıralama yapılır.

5) *En Küçük Benzerlik Değeri*: Elde edilen sonuç listelerindeki aynı resimlerin benzerlik değerlerinden küçük olan alınarak tekrar sıralama yapılır.

IV. DENEYLER VE SONUÇLAR

Bu çalışmada yapılan deneylerle kullanılan resim veritabanına göre hangi benzerlik ölçütünün kullanılabilceğini, ilgi noktası ve öznelilik çıkarma sürecinde birden fazla yöntemin birleştirilmesinin etkisini, çoklu sorgunun tekli sorguya göre etkisini ve değişik sorgu birleştirme yöntemlerinin etkilerini görebilmekteyiz. Şekil 2 'de örnek sorgu resimleri, sonuçları ve çoklu sorgu birleştirme yöntemlerinin örnek sonuçları görülmektedir. Sonuç listeleri 10 nesne içermekte ve nesnelere benzerliklerine göre soldan sağa; sol üst köşedeki birinci, sağ alt köşedeki sonuncu olacak şekilde sıralanmıştır. Şekiller 2 a) ve b) örnek sorgu resimlerini ve arama sonuç listelerini göstermektedir. Şekil 2 c) elde edilen iki sonuç listesinin en iyi sonucu veren sonuç listesi birleştirme yöntemi (*Benzerlik Değeri Toplamı*), Şekil 2 d) ise en iyi histogram aşaması birleştirme yöntemi (*Histogram Ortalama Ağırlıklı*) kullanılarak elde edilen sonuçları göstermektedir.



Şekil 2. Örnek sorgu resimleri ve sonuçları (a) (b), sonuç listesi birleştirme yöntemi (c) ve histogram aşaması birleştirme yöntemi (d)

Önerdiğimiz yöntemi Shen ve ekibinin yöntemi [4] ile aynı resim veritabanı (Caltech256) [6] ve sorgu resimleri kullanarak karşılaştırdık. Caltech256 veritabanında 868 nesne resmi ve 60 sorgu resmi bulunmaktadır. Önerilen çoklu sorgu sisteminin sonuçlarını bu çalışma ile

karşılaştırarak çoklu sorgu yönteminin etkileri gösterilecektir. Shen ve ekibinin asıl amacı sorgu resimlerinden arka planı çıkarmanın resim arama sonuçları etkisini araştırmak olduğundan hem arka planı olan, hem de arka planı temizlenmiş sorgu resimleri kullanılmıştır. Bu nedenle karşılaştırmaların tutarlılığı açısından bu çalışmada her iki sorgu tipi de kullanılmıştır. Deney boyunca sistemin başarısı ortalama duyarlılık (average precision) değerine göre ölçülmüştür. Bütün grafiklerde x eksenini arama sonrası sonuç listesinde gösterilen resim sayısını, y eksenini ise ortalama duyarlılık değerini göstermektedir.

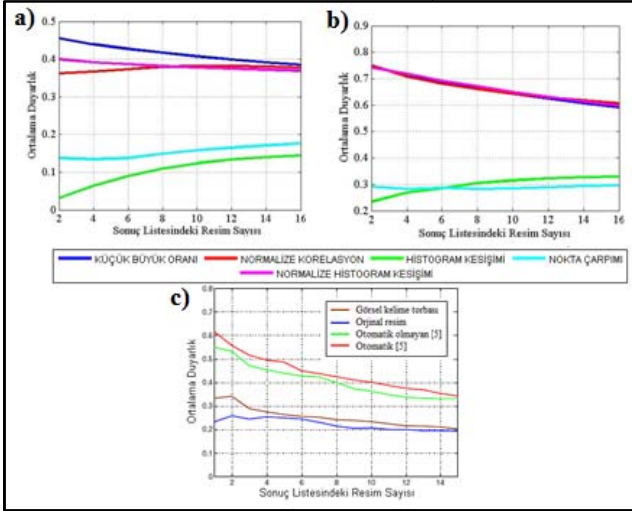
Şekiller 3 a) ve 3 b)'de görüldüğü gibi her iki tip sorgu resimlerinde en iyi sonucu *Küçük Büyük Oranı* yöntemi vermektedir. Bu nedenle bütün deneylerde bu benzerlik ölçütü kullanılmıştır. Şekil 3 c) Shen ve ekibinin yönteminin [5] sonuçlarını göstermektedir. Şekil 3 c)'de kahverengi çizgi bizim de kullandığımız görsel kelime torbası (bag-of-words) yöntemini ifade etmektedir. Şekiller 3 a) ve 3 c) karşılaştırıldığında, aynı resim veritabanı ve aynı sorgu resimleri arka planı çıkarılmadan kullanıldığında Şekil 3 a)'daki en iyi sonuç (*Küçük Büyük Oranı*) Şekil 3 c)'deki görsel kelime torbası yönteminden 0.15 duyarlılık daha iyidir. Geri kalan tüm değişkenler aynı olduğundan bu iyileşmenin sebebi kullanılan ilgi noktası çıkarma yöntemidir. Çalışmamızda Harris ve Surf ilgi noktası bulma yöntemleri birleştirildiğinden klasik görsel kelime yönteminden daha iyi sonuçlar alınmıştır.

Şekil 3 b)'deki *Küçük Büyük Oranı* yöntemi ile elde edilen sonuçlar Şekil 3 c)'deki arka planı çıkarılmış sorgu resimleri kullanılarak elde edilen sonuçlardan (otomatik ve otomatik olmayan) yaklaşık 0.25 duyarlılık daha iyidir. Bu iki yöntem resimlerin arka planlarının Shen ve ekibinin önerdiği otomatik ve otomatik olmayan çıkarılmasının etkilerini göstermektedir [5]. Bizim önerdiğimiz elle arka plan çıkarma yöntemi (Şekil 3 b), Shen ve ekibinin önerdiği her iki yöntemden de (Şekil 3 c) daha iyi sonuç vermektedir.

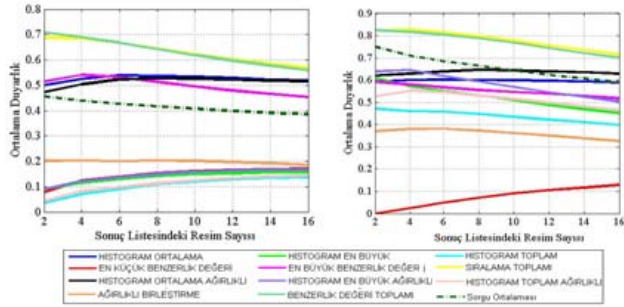
Şekil 4'de ise çoklu sorgu yönteminin tekli sorgu yöntemiyle karşılaştırılması ve sorgu birleştirme yöntemlerinin sonuçları görülmektedir. Sorgu ortalaması tek sorgu resmi yönteminde elde edilen sonucu göstermektedir. Şekiller incelendiği zaman, sonuç listeleri elde edildikten sonra birleştirme yöntemlerinden benzerlik değeri toplamı; histogram aşamasında ise histogram ortalama ve histogram ortalama ağırlıklı yöntemleri kullanıldığında, tekli sorgu yöntemine göre daha iyi sonuç elde edildiği görülmektedir. Arka planı çıkarılmamış sorgu resimlerinde elde edilen artış yaklaşık 0.25 duyarlılık seviyesindeyken, arka planı çıkarılan sorgu resimlerinde artış yaklaşık 0.1 duyarlılık seviyesindedir. Bu sonuçlar çoklu sorgu yönteminin tekli sorgu yöntemine göre daha iyi sonuç verebildiğini göstermektedir.

Şekil 4'den çıkarılabilecek diğer bir sonuç ise histogram birleştirme yöntemlerinde veritabanındaki

histogramlardan elde edilen ağırlıklar kullanıldığında duyarlık açısından daha iyi sonuçlar elde edilmektedir.



Şekil 3. Arka plan çıkarılmamış (a) ve çıkarılmış (b) sorgu resimlerinin ve Shen ve ekibinin [4] çalışmasının sonuçları (c)



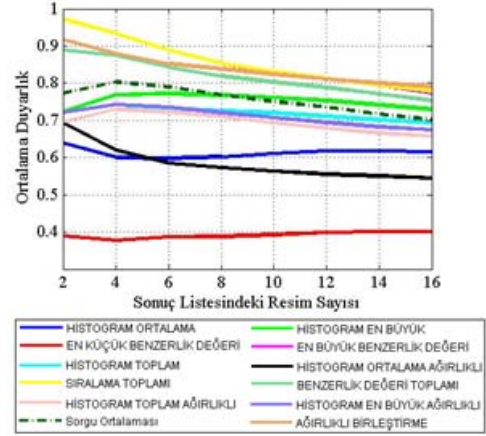
Şekil 4. Çoklu sorgu birleştirme yöntemlerinin arka planı çıkarılmamış (solda) ve çıkarılmış (sağda) sorgu resimlerine göre sonuçları

Şekil 5'de ise sorgu resmi olarak cep telefonu ile çekilmiş resimler kullanılmıştır. En yüksek ortalama duyarlık seviyesi bu resimler kullanılarak elde edilmiştir. Bunun sebebi ise cep telefonu ile bir nesnenin birden farklı açıdan resminin çekilmesidir. Daha önceki sorgu resimlerinde çoklu sorgu için aynı tür nesnenin farklı resimleri kullanılmıştır. Örneğin, çanta kategorisi için farklı çanta resimleri çoklu sorgu için kullanılmıştır. Şekil 5'deki sonuçlarda ise aynı çanta nesnesinin farklı açılardan resmi çekilerek çoklu sorgu oluşturulmuştur. Bu durumda nesneden elde edilen bilgi daha da artırılarak elde edilen ortalama duyarlık seviyesi artırılmıştır. Önerilen yaklaşım mobil ortamda çalışmaya uygundur.

V. SONUÇ

Önerilen yaklaşımın tekli sorgu görüntüsü kullanıldığında klasik görsel kelimeler yöntemine göre, kullanılan birleştirilmiş öznitelik yönteminden dolayı, daha iyi sonuç verdiği gözlemlenmiştir. Shen ve ekibinin yöntemi [5] ile yapılan karşılaştırılmalı sonuçlar, önerilen resim arama yaklaşımının sorgu resmine benzer resim bulma açısından

daha iyi sonuç verdiğini göstermektedir. Deneylerde çoklu resim sorgulama yönteminin tek resimle sorgu yöntemine göre daha iyi sonuç verdiği görülmektedir. Cep telefonu ile çekilen resimlerin sorgu resmi olarak kullanıldığı deneyler, çoklu sorgu yönteminin mobil ortamda kullanılmaya daha uygun olduğunu göstermektedir. Çoklu sorgu yönteminde, histogram aşamasında birleştirme yöntemleri için veritabanındaki resimlerin histogramlarından elde edilen ağırlıkların kullanılması daha iyi sonuçlar vermektedir. Bu çalışmadaki kazanımlar şunlardır:



Şekil 5. Cep telefonu ile çekilmiş sorgu resimlerinin tekli ve çoklu sorgu yöntemleri ile verdiği sonuçlar

- Shen ve ekibinin [5] önerdiği resim arama yönteminden daha iyi çalışan bir yöntem önerilmiştir.
- Çoklu sorgu yöntemi ile tek sorgu yöntemine göre daha iyi sonuç elde edilmiştir.
- Çoklu sorgu yöntemi kullanılarak mobil ortamlar için resim arama sistemi geliştirilmiştir. Cep telefonu ile çekilen resimlerle elde edilen sonuçların en iyi sonuçlar olması çoklu sorgu yönteminin mobil ortam için uygun olduğunu göstermektedir.
- Resim veritabanını göz önünde bulunduran çoklu sorgu birleştirme yöntemleri önerilmiştir.

KAYNAKÇA

- [1] Lee, C. and Lin, M., "A multi-query strategy for content-based image retrieval", *Int. J. of Adv. Inf. Tech.*, 5(2):266-275, 2011.
- [2] Mazloom, M., Habibiyan, A. and Snoek, C., "Querying for video events by semantic signatures from few examples", in *Proc. of the ACM Int. Conf. on Multimedia Retrieval*, pp. 609-612, 2013.
- [3] Arandjelovic, R. and Zisserman, A., "Multiple queries for large scale specific object retrieval", in *Proc. of British Machine Vision Conf.*, pp. 1-11, 2012.
- [4] Lampert, C., "Detecting objects in large image collections and videos by efficient subimage retrieval", in *Proc. of IEEE 12th Int. Conf. on Computer Vision*, pp. 987-994, 2009.
- [5] Shen, X., Lin, Z., Brandt, J., and Wu, Y., "Mobile product image search by automatic query extraction", in *Proc. of European Conf. on Computer Vision, LNCS*, Vol. 7575, pp. 114-127, 2012.
- [6] Griffin, G., Holub, A., Perona, P., "Caltech-256 object category dataset", *Tech. Rep. 7694*, California Inst. of Tech., 2007.