

Yüz ve İsim İlişkisi kullanarak Haberlerdeki Kişilerin Bulunması Finding Faces in News Photos Using Both Face and Name Information

Derya Ozkan, Pınar Duygulu

Bilgisayar Mühendisliği Bölümü, Bilkent Üniversitesi, 06800, Ankara

deryao@cs.bilkent.edu.tr, duygulu@cs.bilkent.edu.tr

Özetçe

Bu çalışmada, haber fotoğraflarından oluşan geniş veri kümelerinde kişilerin sorgulanmasını sağlayan bir yöntem sunulmuştur. Yöntem isim ve yüzlerin ilişkilendirilmesine dayanmaktadır. Haber başlığında kişinin ismi geçiyor ise fotoğrafta da o kişinin yüzünün bulunacağı varsayımıyla, ilk olarak sorgulanan isim ile ilişkilendirilmiş, fotoğraflardaki tüm yüzler seçilir. Bu yüzler arasında sorgu kişisine ait farklı koşullar, poz ve zamanlarda çekilmiş pek çok resmin yanında, haberde ismi geçen başka kişilere ait yüzler ya da kullanılan yüz bulma yönteminin hatasından kaynaklanan yüz olmayan resimler de bulunabilir. Yine de, çoğu zaman, sorgu kişisine ait resimler daha çok olup, bu resimler birbirine diğerlerine olduğundan daha çok benzeyeceklerdir. Bu nedenle, yüzler arasındaki benzerlikler çizgesel olarak betimlendiğinde, birbirine en çok benzeyen yüzler bu çizgede en yoğun bileşen olacaktır. Bu çalışmada, sorgu ismiyle ilişkilendirilmiş, yüzler arasında birbirine en çok benzeyen alt kümeyi bulan, çizgeye dayalı bir yöntem sunulmaktadır.

Abstract

We propose a method to associate names and faces for querying people in large news photo collections. On the assumption that a person's face is likely to appear when his/her name is mentioned in the caption, first all the faces associated with the query name are selected. Among these faces, there could be many faces corresponding to the queried person in different conditions, poses and times, but there could also be other faces corresponding to other people in the caption or some non-face images due to the errors in the face detection method used. However, in most cases, the number of corresponding faces of the queried person will be large, and these faces will be more similar to each other than to others. When the similarities of faces are represented in a graph structure, the set of most similar faces will be the densest component in the graph. In this study, we propose a graph-based method to find the most similar subset among the set of possible faces associated with the query name, where the most similar subset is likely to correspond to the faces of the queried person.

1. Giriş

İnternet üzerindeki haber resimleri zengin bilgi kaynakları olup onların etkin ulaşımı önemlidir. Haberler genellikle kişiler hakkında olmaktadır; bu nedenle kişilere ait sorgulamalar önem kazanmaktadır. Bir kişiye ait bilgileri sorgulamanın alışılmışı yolu o kişinin isminin haber metninde geçip geçmediğine bakmaktır. Fakat bu yol yanlış sonuçlar vermeye eğilimlidir. Bir kişiye ait doğru resimlere erişebilmek için görsel bilgi kullanılmalı ve kişinin yüzü tanınmalıdır. Oysa yüz tanıma henüz çözülememiş bir problem olmaya devam etmektedir.

Son zamanlarda birden fazla veri çeşidi kullanmanın veri erişimi ve analizinde daha iyi sonuçlar verdiği, ve resim ve nesnelere özdevimli olarak sınıflandırmaya olanak sağladığı görülmüştür [1, 8]. Metin bilgisi görsel bilgi ile birlikte sunulduğunda,

yüz tanıma problemi de kolaylaştırılabilir; böylece problem isim ve yüzler arasındaki ilişkiyi bulma problemine dönüştürülebilir [7, 2, 3]. [9 5]' de görüldüğü üzere, metin ve görsel bilginin birlikte kullanılması, tanımaya gerek kalmadan, erişimde daha iyi başarımlar sağlamıştır.

Tüm bunlar göz önünde tutularak, bu çalışmada metin ve görsel bilgiyi birlikte kullanarak sorgu kişisine ait doğru yüzlere erişmeyi sağlayan bir yöntem sunulmuştur. Deneylerde kullanılan veri kümesi, [3] İnternet üzerinden toplanan haber resimlerinden oluşmaktadır ve varolan diğer yüz veri kümelerinden epey farklıdır (Bakınız Şekil 1). Yahoo! News sayfalarındaki haber resimlerinden toplanmış ve haber başlıklarıyla ilişkilendirilmiş çok sayıda resimden oluşan bu veri kümesinde resimler sınırlı ve denetimli ortamlarda değil, gerçek koşullarda çekilmiştir. Bu nedenle de resimler geniş bir çeşitlilikte poz, ışıklandırma ve yüzsel ifade sergilemektedir. Çevresel koşul, yüzün görünmesini engelleyen etkenler, kıyafet, ve yaştaki bu geniş çeşitlilik, veri kümesinin tanınmasını daha zor hale getirmektedir.



Şekil 1. Kullanılan haber resimlerinden örnekler.

Sunduğumuz yöntem; her ne kadar haberde geçen ilgili başka kişilerin yüzleri veya kullanılan yüz çıkarma yönteminin hatasından kaynaklanan yüz olmayan resimler de içerirse (Bakınız Şekil 2); bir isme ilişkilendirilmiş resimlerde o kişinin yüzünün çok daha fazla sayıda bulunacağını varsaymaktadır. Bir diğer varsayım ise, aynı kişiye ait yüz suretlerinin birbirine diğerlerine olduğundan daha çok benzeyeceğidir. Bu varsayımlara dayanarak, önce sorgu ismine ait resimlerin benzerliklerine göre çizgesel olarak betimlenmesi ve ardından bu çizgede en yoğun bileşenin, yani sorgu kişisine ait yüzlerin, çıkarılmasına dayalı bir yöntem sunulmaktadır.



Şekil 2. 'George Bush' ismiyle ilişkilendirilmiş resimler.

Sunulan yöntem, genel yüz tanıma problemine bir çözüm olmamakla birlikte; ismin ve yüzün birlikte geçtiği geniş veri kümelerinde sorgu kişisine ait yüz resimlerine erişim başarımını yükseltmektedir. Bildirinin kalanı şu şekilde organize edilmiştir: Bölüm 2'de isim ve yüzler arasındaki bütünleşmeye yer verilmektedir. Bölüm 3'de yüzler arasındaki

benzerliklerin çizgesel olarak nasıl betimlendiği; Bölüm 4 de ise bu betimlemedeki en yoğun bölgenin bulunmasını sağlayan yöntem anlatılmaktadır. Sunulan yöntem ile gerçekleştirilmiş deneyler sonuçları ise Bölüm 5 verilmektedir.

2. İsim ve Yüzler Arası Bütünleşme

Kullanılan veri kümesi haber resimleri ve başlıklarından oluşmaktadır. Resimde birden fazla yüz ve başlıkta birden fazla isim bulunabileceğinden, hangi ismin hangi yüze ait olduğu tam olarak bilinmemektedir. Fakat şunu söyleyebiliriz ki, bir kişinin isminin geçtiği yerlerdeki resimlerde o kişiye ait yüzün bulunması büyük bir olasılıktır. Bu olasılık göz önünde bulundurularak, ilk aşamada sorgu kişinin isminin geçtiği yerlerdeki yüzler çıkarılmaktadır. Böylece bundan sonra üzerinde arama yapılacak olan veri alanı küçültülmüş olup, elde edilen yeni alanda sorgu kişisine ait yüzlerin çoğunlukla bulunma olasılığı da fazladır.

Her ne kadar elde edilen yeni alanda sorgu kişisine ait yüzler çoğunlukla bulunsun da, bu alan halen o kişiye ait olmayan yüzleri de içermektedir. Sonraki bölümlerde görsel bilginin de kullanılarak, sadece metne dayalı sonuçların nasıl iyileştirilebileceği anlatılacaktır.

3. Yüzler Arası Benzerliklerin Betimlenmesi

Yüzler arasındaki benzerlik bulmada, daha önce Lowe tarafında nesne tanımda kullanılmış ve başarılı sonuçlar elde edilmiş olan SIFT tanımlayıcılar kullanılmıştır [6]. Bu amaçla yüzler bu tanımlayıcı tarafından çıkarılmış ilgi noktaları ile betimlenmiş ve iki yüzün karşılaştırılması esnasında, bir yüzdeki herhangi bir ilgi noktası için diğer yüzdeki en az Öklid uzaklığına sahip noktanın bu noktaya eş nokta olduğu varsayılmıştır. Fakat bu varsayım doğru eşlemelerle birlikte pek çok yanlış eşlemeyi de beraberinde getirmiştir (Bakınız Şekil 3). Yanlış eşlemeleri ortadan kaldırmak amacıyla iki adet kısıtlama uygulanmıştır: geometrik kısıtlamalar, benzersiz eşleme kısıtlamaları.

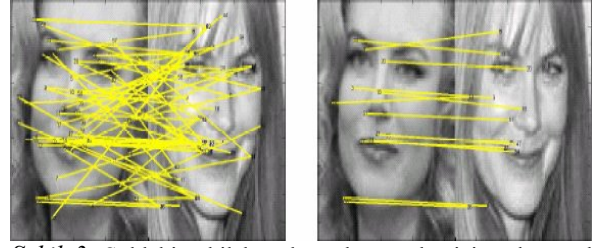
Geometrik Kısıtlamalar

Birbiri ile eşleşmiş iki ilgi noktasının yüzler üzerinde birbirine yakın yerlerde bulunması beklenir. Örneğin, farklı pozlarda bile sol göz genellikle resmin sol ortalarında bulunur. Bu varsayım, eşlenen noktaların yüz üzerindeki göreceli konumları göz önüne alındığında bu noktaların birbirine yakın bağlılık içinde olacağını farz etmektedir.

Göreceli konum olarak birbirinden uzakta bulunan yanlış eşlemeleri ortadan kaldırmak için geometrik kısıtlamalar kullanılmaktadır. Bu amaçla veri kümesindeki rasgele 10 kişi için beşer yüz seçilmiş ve bunlar arasındaki doğru ve yanlış eşlemeler elle işaretlenmiştir. Bu işaretli eşlemeler Bayes sınıflandırıcısı için öğretici olarak kullanılmış ve elde edilen bu sınıflandırıcı bundan sonraki eşlemelerin geometrik uzaklıklarına bakarak bir eşlemenin doğru ya da yanlış olduğuna karar vermekte kullanılmıştır.

Şekil 3'de bir kişiye ait iki yüzün karşılaştırılması ve bu karşılaştırmada geometrik kısıtlamalar uygulanmadan önce ve uygulandıktan sonra eşleşmiş noktalar gösterilmektedir. Görüldüğü üzere yanlış

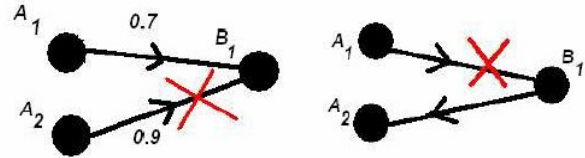
eşlemelerin çoğu ortadan kaldırılmıştır.



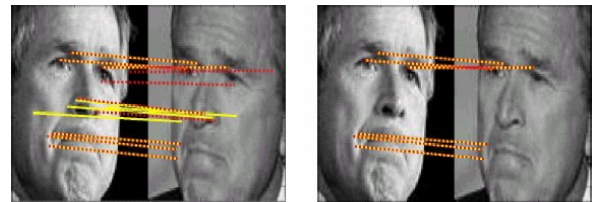
Şekil 3. Soldaki şekil karşılaştırılan yüzler için çıkmış olan tüm ilgi noktalarını ve en az uzaklığa dayalı eşlemelerini göstermektedir. Sağdaki şekilde ise geometrik kısıtlamalar uygulandıktan sonra arta kalan doğru atanmış eşlemeler görülmektedir.

Benzersiz Eşleme Kısıtlamaları

Geometrik kısıtlamalar sonrasında halen bazı yanlış eşlemeler bulunabilmektedir. Bu yanlış eşlemeler çoğunlukla çoklu-eşleme ya da tek-yönlü eşlemelerden kaynaklanmaktadır. Çoklu-eşlemeler, bir noktaya diğer yüzdeki birden fazla noktanın eşlenmesiyle oluşurken; tek-yönlü eşlemeler diğer yüzdeki bir noktaya eşlenen bir noktanın diğer yüzdeki nokta için eş olmamasından kaynaklanmaktadır. Bu yanlış eşlemeler benzersiz eşleme kısıtlamaları ile yok edilerek örneğin A yüzünden B yüzüne eşlenen her noktanın aynı zamanda B yüzünden A yüzüne eşlenen aynı nokta ile eşlenmesi garanti edilmektedir. Örnek bir kısıtlama için Şekil 4'e; kısıtlamalardan önce ve sonraki eşlemeler için Şekil 5'e bakınız.



Şekil 4. A ve B yüzlerinin karşılaştırılmasında A_1 ve A_2 'yi A yüzü üzerinde; B_1 'i de B yüzü üzerinde birer nokta olarak düşünülürse, soldaki resimde çoklu-eşleme, sağdakinde ise tek-yönlü eşleme örnekleri görülmektedir. Çoklu-eşlemedeki eşlemelerden uzaklığı daha fazla olan kaldırılırken (A_2 - B_1 eşlemesi), tek-yönlü eşlemede ise her iki taraftan da eşlenmeyen eşleme kaldırılmaktadır (A_1 ile B_1 eşlenmesine rağmen, B_1 A_1 ile eşlenmemiş).

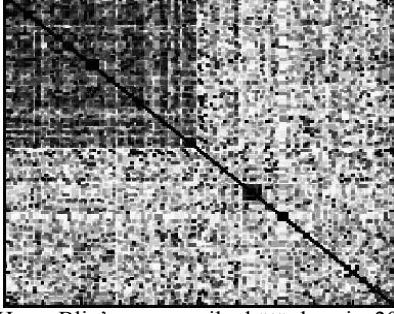


Şekil 5. Solda benzersiz eşleme kısıtlamalarından önce doğru olarak atanmış eşlemeler; sağda ise kısıtlamaların kullanılmasından sonra geri kalan doğru eşlemeler görülmektedir. Kırmızı çizgiler soldaki yüzden sağdakine olan eşlemeler, sarı çizgiler ise sağdaki yüzden soldakine olanları göstermektedir.

Benzerlik Çizgesinin Oluşturulması

Bir sonraki aşamada, kısıtlamalar uygulandıktan sonra elimizde kalan doğru eşlemeler kullanılarak yüzler arasında bir benzerlik çizgesi oluşturulmaktadır. İki yüz

arasındaki uzaklık, bu yüzler arasındaki eşlemelerin ortalama uzaklığı olarak tanımlanır. Daha sonra bu uzaklıklar kullanılarak, isim ve yüz arası bütünleşme ile elde edilen arama alanındaki yüzler arasındaki benzerlikler çizgesel olarak betimlenmektedir. Bu çizge Şekil 6'da görüldüğü gibi matris olarak gösterildiğinde bu matrisin köşegenlerindeki değerlerin 0, ve matrisin de bakışımı olduğu gözlenmektedir.



Şekil 6. 'Hans Blix' sorgusu ile bütünleşmiş 201 resimlik arama alanındaki yüzlerin benzerliklerinin matrisi olarak gösterilmesi. Bu yüzlerden 98 tanesi gerçekten sorgu kişisine ait olup, görsel kolaylık için bu yüzler matrisde sol üst köşeye konmuştur. Koyu renkler küçük uzaklıkları, yani birbirine daha çok benzemeyi ifade etmektedir.

4. Çizgede En Yoğun Bileşenin Bulunması

Yüzler arasındaki benzerlikleri temsil eden çizgede, yüzler düğümlere, aralarındaki uzaklıklar da kırışlere karşılık gelmektedir. Burada amaç bu çizgede en yoğun bileşkeyi bulmaktır. Çünkü en yoğun bileşke bize birbirine diğerlerine olduğundan daha yakın yüzleri verecektir ki bu da bizim varsayımımızla sorgu kişisine ait yüzler olacaktır.

[4]'de bir çizgenin alt kümesi olan S'in yoğunluğu şöyle tanımlanmıştır:

$$f(S) = \frac{|E(S)|}{|S|},$$

$E(S) = \{i, j \in E : i \in S, j \in S\}$ ve E çizgede tanımlı tüm kırışlere karşılık gelmektedir. Burada amaç en büyük f(S) değerine sahip altküme S'yi bulmaktır. Çünkü bu S en yoğun noktaya karşılık gelecektir.

Bu amaçla, önce tüm çizgeden başlanalar f(S) değeri bulunur. Daha sonra her döngüde en az kırış derecesine sahip düğüm atılarak yeni oluşan alt kümenin f(S) değeri hesaplanır. Bu döngü altkümede hiçbir düğüm kalmayana kadar devam eder. Döngü sırasında oluşan tüm alt kümelerden en yüksek f(S) değerine sahip altküme en yoğun bileşen olarak atanır.

Yukarıda anlatılan yöntem en iyi ikili çizgelerde çalıştığından, yöntemi uygulamadan önce elimizdeki benzerlik çizgesi ikili çizgeye dönüştürülür. Bunu yapmak için düğümler arasındaki kırışlere bir eşik değeri konur. Eşik değerinden büyük uzaklık değerlerine sahip kırışler silinerek, küçük olanlara 1 değeri atanır. Böylece elde edilen yeni çizgede 1 iki düğüm arasında bir kırış olduğunu, 0 ise olmadığını ifade eder.

5. Deneyler

Deneylerde kullanılan ve Berg ve ekibi [2] tarafından Internet üzerinden toplanan haber resimleri aslında Yahoo! News'de yayımlanmış milyonlarca resimden oluşmaktadır. Bu resimler üzerine yüz bulma yöntemi

uygulanarak toplam resim sayısı 30.281'e indirilmiştir. Verideki her resim bir isim kümesi ile eşleşmiş olup bu eşlemelerde toplam 13.292 isim kullanılmıştır. Fakat bu isimlerin yarıdan fazlası (9.609 tanesi) sadece bir veya iki kez kullanılmıştır. Aynı zamanda aynı kişiyi tanımlamada birden fazla isim kullanılmış olabilmektedir. Örneğin *George W. Bush* kişisi için şu isimler kullanılmıştır: *George W(1485)*; *W. Bush (1462)*; *George W. Bush (1454)*; *President George W (1443)*; *President Bush (905)*; *U.S. President(722)*; *President George Bush (44)*; *President Bushs(2)*; *President George W Bush (2)*; *George W Bush (2)*. Deneylerde, ismi 200'ün üzerinde ve en çok geçen ilk 23 kişiyi kullanmakta olup bir kişi için varsa farklı isimlerle eşleşmiş resimlerinin hepsini göz önünde bulundurulmaktadır.

Öncelikle, her kişi için isim yüz eşlemesi ile bulunan arama alanında SIFT tanımlayıcılar kullanılarak yüzler arasındaki benzerlikler bulunmuştur. Ardından, daha önce anlatılmış olan kısıtlamalar uygulanmış ve sonuçta gerçekte doğru olan eşlemelerin %72 'si doğru eşleme olarak bulunmuştur. Yöntem ile doğru olduğu söylenen eşlemelerin %71 'i ise gerçekten doğru eşlemelere karşılık gelmektedir.

Sistemin başarımı, benzerlik çizgesini ikili çizgeye çevirirken kullanılan eşik değerine göre değişmektedir. Şekil 8 de 0.55 ve 0.65 arasındaki eşik değerleri için kesinlik-geri getirme değerleri gösterilmiştir. Bu grafiğe bakılarak, 0.575 eşik değeri olarak seçilmiş ve bu eşik değeri için elde edilmiş kesinlik-geri getirme değerleri Şekil 8'de sunulmuştur. Sadece metin kullanıldığında ortalama kesinlik değeri 0.48 iken, bu değer sunulan yöntem ile 0.575 eşik için 0.68'e yükseltilmiş; hatta bazı kişiler için %100 kesinlik değeri elde edilmiştir (*John Ashcroft*, *Hugo Chavez*, *Jiang Zemin*, *Abdullah Gul*). Yine, bu eşik değeri için ortalama geri getirme değeri 0.71 olarak kaydedilmiştir.

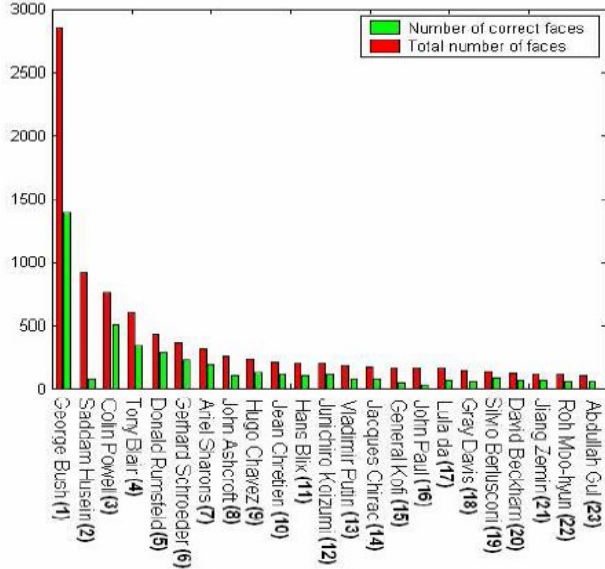
6. Özet ve Sonuçlar

Bu çalışmada, haber metinleri ile ilişkilendirilmiş geniş haber resimlerinde kişi sorgulamayı sağlayan çizgesel bir yöntem sunulmuştur. Bir veri kümesindeki resimler arası benzerlikler betimlenerek, problem bir çizgede en yoğun bileşeni bulma problemine dönüştürülmüştür. Yüzler arasındaki benzerlikleri betimlemek içinse SIFT tanımlayıcılar [6] kullanılarak, iki yüzün birbirine olan uzaklığı birbiri ile doğru eşlendiği varsayılan ilgi noktalarının ortalama uzaklığı olarak atanmıştır.

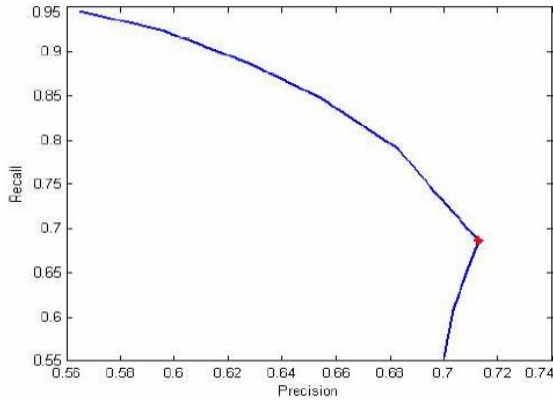
Geniş ve gerçekçi veri kaynakları için yüz tanıma ve sorgulama halen çok zor ve hataya açık bir problemdir. Bu çalışmada, böyle veri kaynaklarında kişi sorgulamaya yarayan bir yöntem sunulmuştur. Sadece metne dayalı yöntemlere oranla kesinlik değerlerinde %20'ye varan iyileştirme sağlanmış ve hatta bazı kişiler için %100 kesinlik, %83.9 geri getirme değerlerine ulaşılmıştır.

4. Bölümde anlatılan çizge yönteminin kullanılmasından önce yüzler arası benzerlik çizgesi ikili çizgeye dönüştürülmekte ve bu nedenle bazı bilgiler kaybedilmektedir. Benzerlik çizgesini ikili çizgeye çevirmeden, yani bazı bilgileri kaybetmeden, yoğun bölgeyi bulan bir yöntemin daha iyi sonuçlar vereceği düşünülmektedir. Gelecekte böyle bir

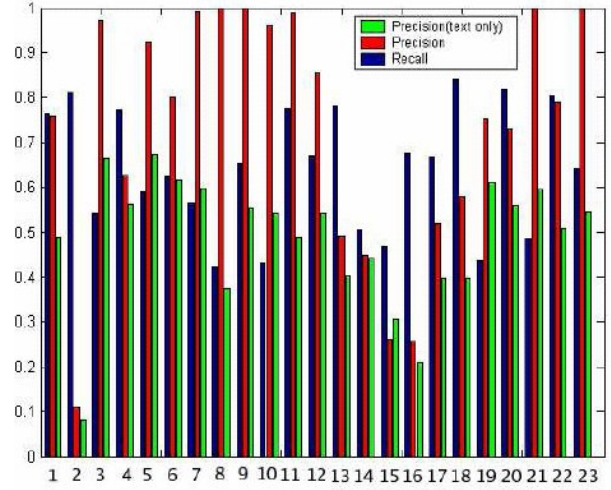
yönteminin oluşturulması üzerinde çalışılmaktadır. Ayrıca yüzler arası benzerlikleri tanımlamada kullanılan SIFT tanımlayıcı yerine başka bir yöntem kullanılabilir ve bunun üzerinde çizge yöntemi uygulanabilir.



Şekil 7. Deneylerde kullanılan 23 kişi. Her kişi için arama alanında bulunan toplam resim sayısı kırmızı, bu alanda gerçekten o kişiye ait resim sayısı ise yeşil çubukla gösterilmiştir.



Şekil 8. Deneyde kullanılan 23 kişi için ortalama kesinlik-geri getirme değerleri grafiği. Kırmızı ile işaretli nokta 0.575 değerine denk gelmekte olup bundan sonraki deneylerde bu değer baz alınmıştır.



Şekil 9. Deneyde kullanılan 23 kişi için 0.575 eşik değerinde kesinlik-geri getirme değerleri grafiği. Mavi çubuklar sunulan yöntemle elde edilen kesinlik, kırmızı çubuklar geri getirme değerlerini göstermektedir. Yeşil çubuklar ise sadece metin kullanılarak elde edilen kesinlik değerlerine aittir.

Teşekkür

Bu çalışma TÜBİTAK Kariyer 104E065 ve TÜBİTAK 104E077 nolu projeler tarafından desteklenmiştir.

Kaynakça

- [1] K. Barnard, P. Duygulu, N. de Freitas, D. A. Forsyth, D. Blei, and M. Jordan. Matching words and pictures. *Journal of Machine Learning Research*, 3, 2003.
- [2] T. Berg, A. C. Berg, J. Edwards, and D. Forsyth. Who is in the picture. In *Neural Information Processing Systems (NIPS)*, 2004.
- [3] T. Berg, A. C. Berg, J. Edwards, M. Maire, R. White, Y.-W. Teh, E. Learned-Miller, and D. Forsyth. Faces and names in the news. In *IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, 2004.
- [4] M. Charikar. Greedy approximation algorithms for finding dense components in a graph. In *APPROX '00: Proc. of the 3rd International Workshop on Approximation Algorithms for Combinatorial Optimization*, London, UK, 2000.
- [5] N. Ikişler and P. Duygulu. Person search made easy. In *The Fourth International Conference on Image and Video Retrieval (CIVR 2005)*, Singapore, 2005. 1
- [6] D. G. Lowe. Distinctive image features from scaleinvariant keypoints. *International Journal of Computer Vision*, 60(2), 2004.
- [7] S. Satoh and T. Kanade. Name-it: Association of face and name in video. In *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1997.
- [8] C. Snoek and M. Worring. Multimodal video indexing: A review of the state-of-the-art. *Multimedia Tools and Applications*, 25(1), 2005.
- [9] J. Yang, M.-Y. Chen, and A. Hauptmann. Finding person x: Correlating names with visual appearances. In *International Conference on Image and Video Retrieval (CIVR '04)*, Dublin City University Ireland, 2004.