

Resim Arama Sonuçlarının Çizge Tabanlı Bir Yöntemle Yeniden Sıralanması

Re-ranking of Image Search Results using a Graph Algorithm

Sare Sevil¹, Hilal Zitouni¹, Nazlı İkizler¹, Derya Özkan², Pınar Duygulu¹

1. Bilgisayar Mühendisliği Bölümü

Bilkent Üniversitesi

{sareg, zitouni, inazli, duygu} @cs.bilkent.edu.tr

2. Institute of Robotics and Intelligent Systems

University of Southern California

derya@usc.edu

Özetçe

Arama motorları aracılığı ile arama yapmak günümüzde internetin en yaygın kullanım amaçlarından biridir. Ancak sorgu sonuçları, özellikle resim arama sorgularında, çoğu zaman sorgu ile ilgili olmayan sonuçlar da içerebilmektedir. Bu çalışma, resimsel aramalarda sorguyla ilgili olmayan yanlış sonuçların belirlenmesini ve tekrar sıralama yöntemleri ile sonuç sıralamasının iyileştirilmesini amaçlamaktadır. Yöntem resimsel benzerliklerin bir çizge ile gösterilmesi ve daha sonra sorguya doğru olarak karşılık gelmesi beklenen en benzer resimlerin çizgedeki en yoğun bileşen olarak bulunması esasına dayanmaktadır.

Abstract

Although one of the most common usages of Internet is searching, especially in image search, the users are not satisfied due to many irrelevant results. In this paper we present a method to identify irrelevant results of image search on the Internet and re-rank the results so that the relevant results will have a higher priority within the list. The proposed method represents the similarity of images in a graph structure, and then finds the densest component in the graph representing the most similar set of images corresponding to the query.

1. Giriş

Günümüz teknolojisinin en gelişmiş örneklerinden biri olan internet her geçen gün kendisini yenileyen niteliklerle yaygın olarak kullanılmaktadır. İnternetin en büyük ve en önemli kullanım amacı ise, 1990'larda geliştirilmeye başlanan arama motorları aracılığı ile, elektronik ortamda araştırma yaparak bilgiye en hızlı şekilde ulaşabilmek olmuştur. Günümüzde yaygın olarak kullanılan arama motorlarından bazıları; Google, Yahoo, Altavista ve MSN olarak sıralanabilir. Bu arama motorları, metin bazlı aramalarda yüksek başarı gösteriyor olmalarına rağmen, resimsel aramalarda ilgili sonuç verme üzerindeki başarı oranları düşüktür. Bunun en önemli sebebi

ise arama motorlarının resim aramada kullanılması halinde arama işleminin yine sözcük bazlı yapılmasıdır. Resimsel aramalarda resmin kendi görsel içeriği yerine resmin yanında yer alan metinlerin ve başlıkların incelenmesi, arama sonuçlarının içerisinde çok sayıda ilgisiz verilerin de yer almasına sebep olmaktadır.

Bu makalede anlatılan yöntem, arama sonuçlarının resmin içeriğine dayalı olarak yapılandırılmasını amaçlamaktadır. Buna göre, yapılan görsel sorgulamalar üzerine çıkan sonuçlarda ilgisiz veriler belirlenmekte ve arama sonuçları bu bilgiler dahilinde tekrar sıralanarak kullanıcı açısından daha kullanışlı ve tercih edilebilir sonuç sıralamaları oluşturulmaktadır. Böylece arama üzerine çıkan sonuçlarda en çok beklenen görsel içerikteki resimlere öncelik sağlanmış olacak ve bu resimler sıralamada baş kısımlarda yer alacaktır.

2. İlgili Çalışmalar

Literatürde, arama motorlarından gelen sonuçların belirli bir yöntemle tekrar sıralanmasını öneren çok fazla çalışma bulunmamaktadır. Az sayıda olan bu çalışmaların içersinden Ben-Haim ve ekibinin çalışmasını, [3], bu makalede önerilen yönteme yaklaşım tarzı açısından en çok yakınlık gösteren çalışma olarak gösterebiliriz. Ben-Haim ve ekibinin [3] çalışmasında izlediği ReSPEC adlı yöntem şu şekilde özetlenebilir: arama sonuçlarının belirli bir orana göre başlarında kalan resimler, öncelikli olarak "blob" adı verilen parçalara bölünür. Bu ayrılan parçalardan öznitelik küme vektörleri çıkarılır ve çıkarılan bu vektörlere "mean shift" algoritması uygulanarak vektörler gruplara ayrılır.

Çalışmada, bu grupların, sonuçların önem sırasıyla ilgili bilgi içerdiği; grupların yoğunluk oranlarının, o grupta bulunan sonuçların önemiyle doğru orantılı oldukları ileri sürülmektedir. Buradan yola çıkarak, sonuçlarda *blob*lara ayrılmadan kalan her bir resim, gruplara olan uzaklıklarına göre, en yakın olduğu gruba eklenir ve yeni grup sonuçları elde edilir. Yoğun grup ve resim önemi ilişkisi doğrultusunda

da yöntemin çıktısı, önem sırasına göre tekrar sıralanmış arama sonuçları olarak sunulur.

Alandaki başka bir önemli çalışma Zisserman ve ekibi tarafından hazırlanmış olan [4] olarak gösterilebilir. Çalışmanın amacı, internette belirli bir konu üzerine araştırma yapılarak resim toplayıp kullanılabilir bir veri kümesi oluşturmaktır. Önerilen yöntem bu işlemi, Google kullanarak hem webden hem de Google'ın resimsel aramalarından gelen sonuçları önem sırasına göre yeniden dizip, önemli resimleri veri kümesine ekleyerek gerçekleştirmektedir.

Burada sunulan yöntemden farklı olarak Zisserman ve ekibinin çalışması tamamen görsel yöntemlere dayanmaktadır. Sunulan yöntemin iki ana adımından ilkinde sözcük temelli özneliktik vektörleri kullanılarak kaba bir sıralama oluşturulduktan sonra ikinci adımda resimsel özneliktik vektörleriyle önem sırasının son haline ulaşmıştır.

Çalışmada, internette toplanarak elde edilmiş olan büyük ve test aşamasında kıyaslama amacıyla kullanılmaya elverişli veri kümesinde yaptıkları yeniden sıralama yöntemi, internet üzerinden araştırma yapılarak belirli kategoriler için veri kümeleri oluşturmayı hedeflemektedir. Çalışmada kullanılan resimler Google'ın hem web hem de resimsel aramalarından gelen sonuçları içermektedir.

Gözlemlere göre, internet üzerinde yapılan resimsel aramalarda yanlış sonuçlar elde edilse de, alınan doğru sonuçlar genellikle çoğunluğu oluşturmaktadır. Bu gözlemden yola çıkarak çalışmamızda, internet yerine, verilen bir veri kümesinde sadece yüz resimleri üzerine uygulanmış başka bir yöntemi kullandık. Çalışmamızın arama sonuçlarındaki ilgisiz verilerin belirlenmesi ile ilgili olan kısmıyla ilişkili bu yöntem, Özkan ve Duygulu'nun çalışmasında anlatılmaktadır [1]. Bu çalışmada yapılan sorgulamalar, yüzler ve yüzlerin etrafında geçen yazılardaki isimleri eşleştirerek, doğru yüzü doğru isime ilişkilendirmeyi hedeflemektedir. Çalışmada, yüzler ve isimler arası yanlış eşleşmelerin elenmesi için çizgede en yoğun bileşenin (densest component) bulunması yöntemi uygulanmıştır. Bir sonraki bölümde daha detaylı bir şekilde anlatılacak olan bu yöntem, Charikar tarafından [2] çalışmasında sunulmuştur; düğümler ve düğümleri bağlayan kenarlardan oluşan, tam bağlı (*fully connected*) bir çizgedeki en yoğun düğüm alt kümesini bulmak için kullanılmaktadır.

Özkan ve Duygulu bu makalelerinde, çizgede en yoğun bileşenin bulunması işlemini uygularken Charikar'ın sunmuş olduğu algoritmayı [2] kullanmışlardır. Özkan ve Duygulu, haberlerde geçen yüzleri ve bu yüzlere ait isimleri ilişkilendiren, ve böylece bu tür veritabanları içinde, kişilerin isimleri ile yapılan sorgulamalardan verimli şekilde sonuç alınan bir yöntem geliştirmişlerdir. Bu yöntemde öncelikli olarak isim ve yüzler arası bütünleşme bulunmuş, bunun üzerine yüzler arasındaki benzerliklerin çizelgesel olarak nasıl betimlendiği çıkarılmıştır. Yöntemde, çıkarılan bu betimlemelerdeki en yoğun bölgenin, en yüksek olasılıkla sorgulanan kişi olduğu tezi öne sürülmüştür. Böylece betimlerdeki en yoğun bölge çizgede en yoğun bileşenin (densest component) bulunması yöntemi kullanılarak belirlenip sorgu sonuçları bulunmuştur.

3. Motivasyon

Bu çalışmadaki amacımız, Özkan ve Duygulu'nun yöntemini genelleyip, hem internet ortamındaki sorgulamalara uygulamak hem de sorgu çeşidi olarak her hangi bir sınırlama

tanımlamak olmuştur. İnternetteki, resimsel sorgulamalarda doğru sonuçların çoğunlukta olacağı gözleminden yola çıkarak çizgede en yoğun bileşeni bulma yöntemi, Özkan ve Duygulu'nun çalışmalarından esinlenerek yöntemimize uygulanmıştır. Çalışmamızın en önemli özelliği kullandığımız yöntemin öğreticili bir yöntem olmamasıdır; öğreticili bir yöntem her sorgu için bir öğreti oluşturmayı gerektireceği için sorgulamaları sınırlandırır. Yöntemimiz tüm sorgu çeşitleri için uygulanabilecek düzeydedir.

4. Yapılan Çalışma

Arama motorları kullanılarak yapılan aramalar sonucunda sorguya en bağlı bilgilerin ilk sıralarda çıkmasını amaçlayan bu yöntem için izlenen yöntemi Şekil-1 deki akım şeması aracılığıyla daha net açıklayabiliriz

Yöntemimizde ilk önce her bir kategori için ağ örün tarayıcısı (*web crawler*) kullanılarak belirli bir sorgunun sonuçlarını oluşturan resimler elde edilmiştir. Daha sonra bu resimlerden ayrılan altküme içerisindeki her bir resmin birbirleri ile arasındaki benzerlik matrisi bulunmuş ve bu matrise dayanarak tam bağlantılı bir benzerlik grafiği oluşturulmuştur. Benzerlik grafiği üzerinde kullanılarak Charikar'ın algoritmasının uygulanması üzerine en yoğun bileşenler bulunmuş ve altküme dışında kalan diğer sonuçlar en yoğun bileşen **ortalamasına** göre benzerlik grafiğine eklenmiştir. Bütün bu işlemler sonucunda elde edilen benzerlik grafiği doğrultusunda ise sonuçlar sorguyla ilişki derecesine göre tekrar sıralanmıştır.

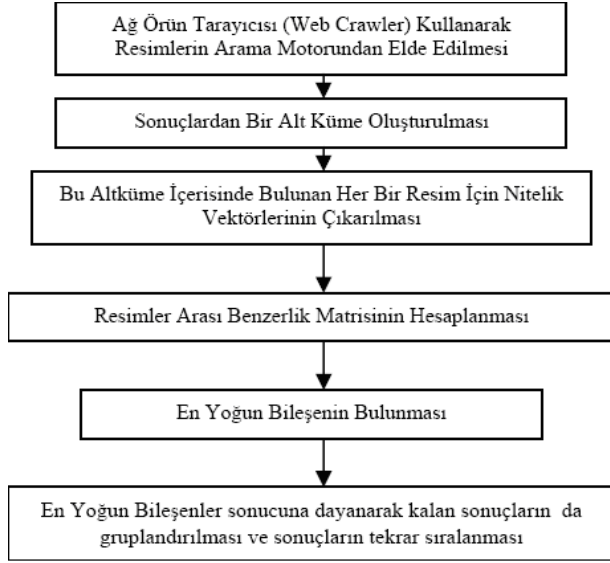
Aşağıda, sorgulamanın sonuçlarını oluşturan resimlerin içerisinden bir altküme çıkarıldıktan sonra yapılan işlemler detaylıca anlatılmaktadır:

Nitelik Vektörlerinin Çıkarılması: Elimizdeki altkümedeki her bir resim için nitelik vektörleri çıkarılmasında renk histogramları kullanılmıştır. Çıkarılan renk histogramları sayısına ise resimler arası benzerlikler hesaplanmıştır.

Benzerlik Matrisinin Oluşturulması: Benzerlik matrisi, verilen bir sorgu sonuçları kümesindeki her resmin diğer resimlere olan uzaklığının tutulduğu matristir. Yöntemimizde iki resim arasındaki bu uzaklık *chi_squared distance* algoritması kullanılarak hesaplanmıştır. Bu algoritma için ise yukarıda bahsetmiş olduğumuz resimlere ait nitelik vektörleri kullanılmıştır.

Benzerlik Grafiğinin Oluşturulması: Boolean bir matriste tutulan benzerlik grafiğinde her düğüm bir resmi, düğümler arası her bir bağlantı ise resimler arasında var olan ilişkiyi temsil etmektedir. Benzerlik grafiğini oluşturan yöntemde farklılık matrisindeki değerler kullanılmıştır. Bu yöntemde göre, farklılık matrisleri incelenerek bir eşik değeri belirlenmiş ve bu eşik değerinin üzerinde kalan değerler, birbirleri ile ilgili olmayan resimler olarak nitelendirilmiştir. Diğer taraftan bu eşik değerinin altında kalan değerler ile benzerlik grafiği için gerekli olan resim çiftleri elde edilmiştir.

Bu tanımdan yola çıkarak farklılık matrisiyle aynı boyutta tanımlanan benzerlik grafiğinin tutulduğu matrisin içindeki değerler farklılık matrisinde eşik değerinin altında kalanlar "2", ve üzerinde kalanlar ise "0" olmak üzere doldurulmuştur. Yeni yaratılan bu matris ise Charikar[2] tarafından sunulmuş olan algoritma kullanılarak uygulanan en yoğun bileşenin bulunması metoduna girdi olarak gönderilmiştir.



Şekil 1: Kullanılan yöntemin akım şeması

En yoğun bileşenin bulunması: Verilen bir benzerlik grafiği için en yoğun bileşen Charikar'ın algoritması[2] kullanılarak bulunmaktadır. Bu algoritmada incelenen grafiğin tüm alt kümeleri için bir yoğunluk değeri hesaplanmaktadır. Tüm bu hesaplanan değerler içinde en büyük yoğunluk değerine sahip alt küme grafikteki en yoğun bileşeni temsil etmektedir.

Verilen bir alt küme S için yoğunluk Charikar tarafından önerilen [2] aşağıdaki formül ile hesaplanmaktadır.

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

Burada:

- $f(S)$: S'nin yoğunluğu,
- $E[S]$: S kümesindeki kenar sayısı,
- $|S|$: S kümesindeki düğüm sayısı

olarak tanımlanmaktadır.

Yöntemimiz, tüm sorgu çeşitlerine uygulanabilir olmakla birlikte, öncelikli olarak doğa sahneleri üzerinde sınanmıştır. Kullandığımız veriler Google'da yapılan 7 farklı sorgunun sonuçlarından oluşmaktadır. Sorgularımızı bina, sahil, dağ, kar, orman, okyanus ve gün batımı olarak sıralayabiliriz. Çalışmamızı basit ve sınırlı sayıda sorgulama üzerinde denememizin sebebi ilk aşamada öncelikli olarak sonuçta bir gelişme kaydedilip edilmediğini görmek istememizden kaynaklanmaktadır. Bu nedenle ancak kategorilerimiz arttıkça ve çeşitlendikçe elde edilen başarı yüzdesi üzerinde daha çok yorum yapabileceğiz.

Yöntemimiz sorgu temelli çalıştırılmıştır. Öncelikli olarak veri kümesi yaratılırken her sorgunun sonuçlarındaki ilk 10 resimden altkümeler oluşturulmuştur. Benzerlik grafiğinin öncelikle bu ilk on resmi içeren altkümelerden çıkarılmasının sebebi ise şöyle açıklanabilir: yöntemimiz yoğunluk oranına göre çalıştığı için istenilen sonuca ulaşabilmek, elimizdeki veri kümesindeki *doğru* resim sayısının *yanlış* resimlere oranla

daha fazla olmasına bağlıdır. Bu nedenle grafik çıkarma yönteminin ilk aşamada veri kümesinin tümü yerine belirli bir altkümesinde çalışılması uygun görülmüştür. İncelememizde Google'da yapılan resimsel aramalarda ilk 10 resimden oluşturulan altkümelerde en etkin sonuçlar elde edilmiştir ve bu sebeple sınır olarak ilk 10 resim belirlenmiştir.

Altküme üzerinde gerekli hesaplamalar yapıldıktan ve geri kalan resimler de buna göre gruplandırıldıktan sonra yöntemimiz her kategorideki ilk 100 resmin sıralanmasındaki başarı yüzdesini gösterecek şekilde çalıştırılmıştır. Bu kıyaslamayı yapabilmek için ise her kategorideki ilk 100 resim için el yordamıyla 1x100 bir lük doğruluk matrisi oluşturulmuştur. "1" ve "0" lardan oluşan doğruluk matrisimizde "1" ler sorgu ile ilgili olan sonuçları, "0" lar ise sorgu ile ilgisiz olarak gelen sonuçları simgelemektedir.

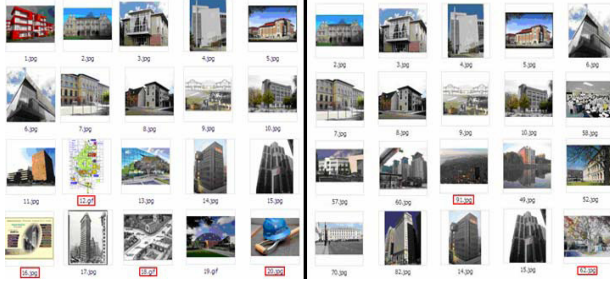
Yeniden sıralama işlemi için ilk önce resimler arası benzerlik ağırlıklarını tutan bir benzerlik matrisi oluşturulmuştur. Bu matriste her kategorideki 100 resmin birbirine olan uzaklıkları tutulmaktadır. Bu sebeple benzerlik matrisimiz $m(i,j)$ 100x100 lük bir matris olup her (i,j) elemanı içerisinde i 'nci resmin I_i , j 'nci resme I_j olan benzerlik ağırlığını tutmaktadır. Benzerlik ağırlığı için ise i 'nci resmin nitelik vektörünün, j 'nci resmin nitelik vektörüne olan uzaklığı *chi_squared* algoritması kullanılarak hesaplanmıştır. Daha sonra bu matris kullanılarak en yoğun bileşenin bulunması işlemi uygulanmıştır. En yoğun bileşenin bulunması için her kategorideki ilk 10 resim alınmış ve bu resimlerin benzerlik matrisleri kullanılmıştır. En yoğun bileşenin bulunması için kullanılan Charikar'ın sunmuş olduğu algoritmaya[2] ise girdi olarak benzerlik matrisi kullanılarak oluşturulmuş olan yeni bir matris verilmiştir. Bu yeni matris "0" ve "2" lardan oluşmuş olup, oluşturulmasında deneyler aracılığıyla (ampirik olarak) 0.45 olarak belirlenmiş olan eşik değeri kullanılmıştır. Bu eşik değerinin altında kalan değerler benzerlik oranı yüksek olduğu için "2", altında kalanlar ise benzerlik oranı düşük olduğu için "0" değerini almışlardır. Bu matris herbir 10 resmin birbirine benzer ya da benzemez olduğunu belirten bir matristir. Bu matris kullanılarak da 1x10 luk bu matrisin satır değerlerinin toplamını tutan bir matris oluşturulmuştur. Buradaki minimum değer bize en ilgisiz resmi verecektir. Bu her yinelemeye(iteration) bu matristeki en küçük değer çıkarılarak yeni *en yoğun bileşenler grubu* bulunur. En yoğun bileşen grupları arasındaki en büyük değer bize ilgisiz düğümlerin sona atılmış olduğu ilgili kategoriye ait resimleri başa alan sıralamayı verir. Daha sonra bu sıralamaya dayanarak veri kümemizde kullandığımız bu 10 resmin dışında kalan resimler yeniden sıralanır.

Sonuçlarımızdan bahsetmeden önce neye bağlı olarak bu sonuçlara ulaştığımız anlatılacaktır: daha önceden de belirtilmiş olduğu üzere başarı yüzdemiz elimizdeki veri kümesinin yeniden sıralanmış halindeki *doğru* resimlerin başlarda olma oranının eskisine göre daha yüksek olmasına bağlıdır. Bu verileri alabilmek için *recall* ve *precision* diye nitelendirdiğimiz iki orandan bahsetmemiz gerekmektedir.

Geri Getirme (recall) olarak adlandırdığımız geri getirme yüzdesi, arama motoru üzerinde yapılan sorgunun verdiği sonuçlar içerisindeki doğruluk yüzdesini gösterirken, *precision* ya da diğer bir deyişle kesinlik yüzdesi, bizim uyguladığımız yöntemle çıkan doğruluk yüzdesinin oranını vermektedir.

Şekil 2'de 7 farklı sorgulama için yapılmış olan arama sonuçlarının ilk ve programımız tarafından yeniden

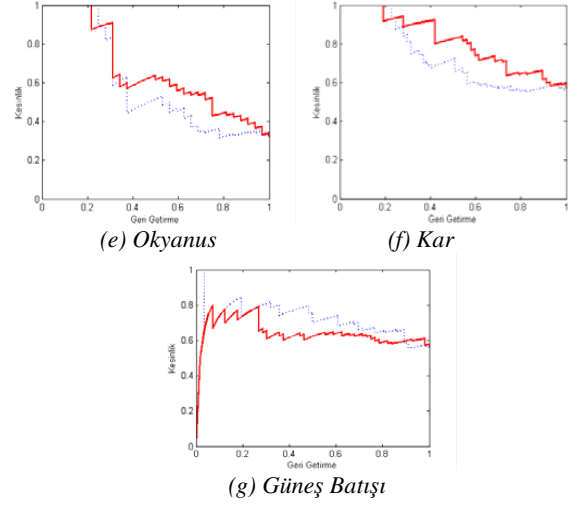
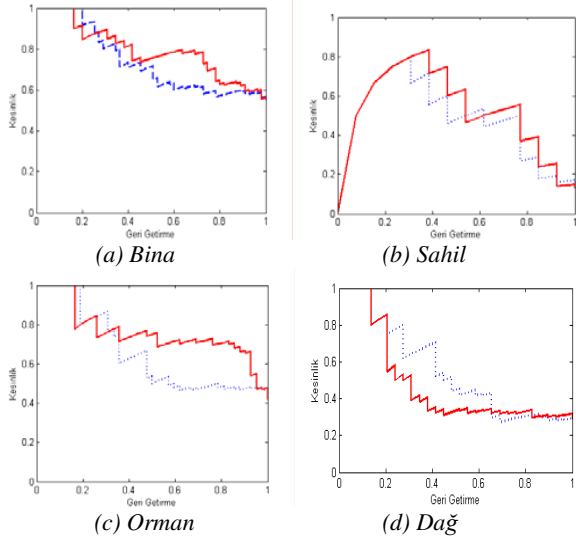
sıralandıktan sonraki durumları için elde edilen kesinlik (precision) – geri getirme (recall) grafikleri gösterilmiştir. Mavi grafik arama sonuçlarının ilk halini gösterirken, kırmızı grafik yeniden sıralanmış sonuçların grafik çizgisini oluşturmuştur.



Şekil 2: Bina sorgusu için çizginin: sol tarafı Google'dan gelen sıralama; sağ tarafı yöntemimizden çıkan sıralama. Yanlış resimlerin isimleri kırmızı kutu içine alınmıştır.

Şekil 3'te verilen Geri Getirme/Keskinlik grafikleri incelendiğinde yöntemimizin genel olarak iyi sonuçlar verdiği görülmektedir. Grafiklerden en kötü olan Güneş Batımı ve Dağ'ın sonuçlarının bu şekilde çıkmasının sebebi bu sorguların sonuçlarında sorguyla ilgisiz resimlerin daha fazla olmasıdır. Dolayısıyla en yoğun bileşen yöntemi başarısız olmuştur. Buna ek olarak bu aşamada sadece basit bir öznitelik olan renk histogramlarının kullanılmış olması da sonuçların kötü çıkmasında etkili olmuştur.

Şekil 2'de ise bina sorgusu için ilk yirmi resmin Google'dan alınan ilk sıralaması ve yöntemimiz sonucu oluşan sıralama gösterilmektedir. İlk yirmi resimde bulunan toplam yanlış resim sayısından da görüldüğü üzere yöntemimiz ilk 20 resmin orjinal sıralamasında yer alan yanlış resim sayısını düşürmüştür.



Şekil 3: Şekilde 7 farklı sorgu sonuçlarına ait Geri Getirim – Keskinlik Grafikleri verilmiştir. İnternet üzerinden yapılan arama sonuçları mavi, nokta nokta çizgilerle, bizim uygulamış olduğumuz yöntemin sonuçları ise kırmızı, düz çizgi ile gösterilmiştir.

5. Sonuç

Bu çalışmada internet üzerinde bir arama motoru aracılığı ile yapılan resim sorgularının sonuçlarını sorguyla ilişkiye göre resimleri önem sırasına dizen bir yöntem geliştirilmiştir. Yöntem öğreticili bir yöntem değildir ve tüm sorgu çeşitlerine uygulanabilecek şekilde tasarlanmıştır. Şu aşamada sadece basit sorgularda denenmiş olan yöntemimizin daha zor sorgularda iyi performans gösterebilmesi için gelecekte daha karmaşık öznitelik vektörleri kullanmayı düşünüyoruz.

6. Teşekkür

Bu çalışma 104E065 ve 104E077 nolu Tübitak projeleri tarafından desteklenmektedir.

7. Kaynakça

- [1] Derya Ozkan, Pinar Duygulu "A Graph Based Approach for Naming Faces in News Photos" CVPR 2006, New York City.
- [2] Moses 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.
- [3] Nadav Ben-Haim, Boris Babenko and Serge Belongie "Improving Web Based Image Search via Content Based Clustering", [IEEE Xplore] SLAM 2006, New York City.
- [4] F. Schroff, A. Criminisi, and A. Zisserman "Harvesting Image Databases from the Web", ICCV 2007.