

İçerik Tabanlı Görüntü Erişim Sistemleri Content-Based Image Retrieval Systems

İrem SOYDAL*, Umut AL** ve Umut SEZEN***

Öz

Gün geçtikçe sayısal görüntü koleksiyonlarının sayısı hızla artmakta ve kullanıcıların aradıkları sayısal nesnelere erişim güçleşmektedir. Gerek bireyler gerekse kurumlar, görüntü arşivlerini yaratırlarken erişimde ciddi sıkıntılar yaşamaktadırlar. Görsel bilgi, kelimelerle tanımlanamayan fakat sezgisel olarak algılanabilen bir yapıya sahiptir. Bu yapı, metin tabanlı bilgiden fazlasını talep etmemize neden olmaktadır ve içeriği değerli kılmaktadır. Aynı yapı bu tip bilgiyi, tanımlanması, depolanması ve dolayısıyla erişilmesi güç hale getirmektedir. Konuyla ilgili olarak özellikle 90'lı yıllarda artan çalışmalar, görsel bilgiye erişim konusunda belli bir noktaya ulaşılmasını sağlamıştır. Yapılan çalışmalarda kesin çözümler üretilmemiş olsa da, önemli bir aşama kaydedildiği açıktır. Bu çalışmada, görüntü erişim sistemleri ile ilgili genel bilgiler verilmekte, içerik tabanlı görüntü erişimi, içerik tabanlı görüntü erişim sistemleri ve söz konusu sistemlerin özellikleri gibi konular ele alınmaktadır.

Anahtar sözcükler: İçerik tabanlı görüntü erişimi, Görüntü erişimi, Görsel bilgi erişim.

Abstract

Digital image collections are expanding day by day, and image retrieval becomes even harder. Both individuals and institutions encounter serious problems when building their image archives and later when retrieving the archived images. Visual information cannot

*Arş. Gör.; Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü Beytepe 06800 Ankara (soydal@hacettepe.edu.tr).

**Arş. Gör.; Hacettepe Üniversitesi Bilgi ve Belge Yönetimi Bölümü Beytepe 06800 Ankara (umutal@hacettepe.edu.tr).

***Dr.; Hacettepe Üniversitesi Elektrik ve Elektronik Mühendisliği Bölümü Beytepe 06800 Ankara (U.Sezen@hacettepe.edu.tr).

be fully expressed in words and normally depends on intuitive human perception. Consequently, this causes us to find the plain text-based information inadequate, and as a result, increases the value of the visual content. However describing, storing and retrieving the visual content is not simple. The research activities in this area, which escalated in the 90's, have brought several solutions to the understanding, design and development of the image retrieval systems. This article reviews the studies on image retrieval systems in general, and content-based image retrieval systems specifically. The article also examines the features of content-based image retrieval systems.

Keywords: Content-based image retrieval, Image retrieval, Visual information retrieval.

Giriş

Bilgi miktarındaki hızlı artışa paralel olarak bilgi erişimle ilgili sorunlar da artmakta ve çeşitlenmektedir. Büyük miktarlardaki bilgi yığınları içinde kullanıcıların istenilen bilgiye erişmesi zorlaşmaktadır. Tonta (2001, s. 200) bir bilim disiplini olarak bilgi erişimin 50 yıllık geçmişi olduğunu ve bilgi erişimin “bilgi toplama, sınıflama, kataloglama, depolama, büyük miktardaki verilerden arama yapma ve bu verilerden istenen bilgiyi üretme (veya gösterme) teknik ve süreci” olarak tanımlandığını ifade etmektedir. Bu sürecin sağlıklı işleyebilmesi için dizinleme ve sınıflama sistemleri oluşturulmakta, arama motorları ve arama robotları tasarlanmaktadır. Bununla birlikte oluşturulan sistemlerin ve tasarlanan araçların yeterli olmadığı görülmektedir (Tonta, 2001, s. 204). Günümüzde metinsel erişimle ilgili problemleri bile tam anlamıyla çözümlenememişken, karşımıza görsel bilgiye erişimdeki sorunlar çıkmaktadır. Sayısal görüntü miktarındaki artış¹ ve kullanıcıların deneyimlerindeki sınırlamalar, görsel bilgi erişim sistemlerinin sorunlarını artırmaktadır.

Görsel bilgi iki temel bileşenden oluşmaktadır: Bunlardan biri, nesne ile ilgili bilgiyi içeren üst veri, diğeri ise görsel nitelikler olarak adlandırabileceğimiz nesnenin kendisinde bulunan bilgidir (Gupta ve Jain, 1997, s. 72). Üst veri metinsel nitelik taşımakla birlikte, genellikle veri tabanları ile ilişkilendirilen bir yapıya sahiptir. Görsel nitelikler ise görsel nesneden bir dizi bilgisayar işlemi sonucunda elde edilmektedir. Söz konusu bilgisayar işlemleri arasında özellikle görüntü işleme ve çeşitli geometrik hesaplamalar önemli yer tutmaktadır (Gupta ve Jain, 1997, s. 72).

Görsel nesnelere görüntü ve hareketli görüntü olmak üzere iki grupta incelenmektedir (Eidenberger, 2004). Elektronik ortamda “.jpeg, .tiff, .gif” gibi

¹Şubat 2005 itibarıyla Google görüntü veri tabanındaki resim sayısının 1,1 milyarı, Yahoo'nun görüntü veri tabanının içerdiği resim sayısının ise 1,5 milyarı aştığı görülmektedir (Price, 2005).

dosya formatlarıyla depolanan fotoğraf, resim, grafik ve benzeri biçimlerdeki materyaller için "görüntü"; .mpeg, .avi, .mov gibi dosya formatları ile depolanan klip ya da video görüntüsü özelliği taşıyan materyalleri tanımlamak için de "hareketli görüntü" ifadeleri kullanılmaktadır. Hareketli görüntüler birbirinin peşi sıra gelen sabit görüntülerden oluşmakla birlikte, sesli olabilme ve belirli bir zaman içerisinde devam etme gibi özellikler onları sabit görüntülerden ayırmaktadır. Çalışmamız kapsamında hareketli görüntüleme erişimle ilgili konulara yer verilmemektedir. Çalışmamızda görüntü erişim sistemlerinin neler olduğu açıklanmakta, içerik tabanlı görüntü erişim, içerik tabanlı görüntü erişim sistemleri ve söz konusu sistemlerin özellikleri ele alınmaktadır. Ayrıca literatüre bakıldığında konuyla ilgili bazı kavramların farklı kaynaklarda farklı şekillerde² ifade edildiği görülmektedir. Kavram karmaşasına engel olmak amacıyla Türkçeleştirdiğimiz ilgili terimlerin orijinal karşılıkları parantez içinde verilmektedir.

Görüntü Erişim Sistemleri

Görüntü erişimi ile ilgili çalışmaların kökeni 1970'lere kadar gitmektedir (Rui, Huang ve Chang, 1999, s. 39). Zaman içinde söz konusu çalışmalara olan ilgideki hızlı artış ve buna paralel olarak konuyla ilgili yayın sayısının çokluğu açıkça gözlenmektedir (Eidenberger, 2004). Görüntü erişimi üzerine yapılan çalışmalar farklı şekillerde karşımıza çıkmaktadır. Bunlar metin tabanlı erişim, içerik tabanlı erişim ve kullanıcı etkileşimli görüntü erişim sistemleridir (Goodrum, 2000, s. 63).

Metin tabanlı görüntü erişiminde sözcük ya da sözcük grupları kullanarak bir görsel malzemeyi her yönüyle ifade edebilmek önemlidir. Ancak, kullanıcının söz konusu materyali farklı yönleriyle öne çıkarabileceği ve bu özellikleri kullanarak bir görsel malzemeyi arayabileceği düşünüldüğünde metin tabanlı görüntü erişim sistemlerinin yetersizliği daha iyi anlaşılmaktadır. Bir tek resmin her yönüyle ifade edilebilmesi için çok fazla dizin terimi gerekmesi ve bunun insanlar tarafından çoğunlukla elci sistemlerle yapılması harcanan emeği çok fazla artırmaktadır. Bunu aşmak ve görüntü tanımlamalarını yapabilmek için metin tabanlı ontolojiler ya da sınıflama şemaları hazırlanmıştır (Cawkell, 2000, s. 53; Goodrum, 2000, s. 63; Goodrum ve Spink, 2001, s. 296; Rui ve diğerleri, 1999, s. 40).

Kullanıcı etkileşimli görüntü erişim sistemleri ile farklı disiplinlerde bulunan, birbirinden farklı bilgi gereksinimlerine sahip kullanıcıların sorgu türleri analiz edilebilmektedir. Aynı zamanda bu tip sistemler kullanıcıların bilgi arama

²Örneğin "feature extraction" için öznelilik bulma, özellik çıkarımı, öznelilik dönüştürme, öznelilik çıkarımının kullanılması gibi.

davranışları ile sistemin örtüşmesini sağlamayı amaçlamaktadır (Goodrum, 2000, s. 65). Burada önemli olan kullanıcının bilgi gereksinimidir ve söz konusu bilgi gereksinimi kullanıcıların arka plan bilgileri ve ilgi alanları ile örtüşmektedir. Çeşitli disiplinlere yönelik olarak o disiplinin kullanıcılarının gereksinim duyduğu görsel bilgiye erişim sağlamayı kolaylaştıracak sistemler tasarlanmaktadır. Örneğin; gazetecilik (Markkula ve Sormunen, 1998; 2000), kriminoloji (Lee ve diğerleri, 2004; Wu ve Narasimhalu, 1994), tıp (Lehmann, Meinzer ve Tolxdorff, 2004; Lehmann ve diğerleri, 2004; Müller, Michoux, Bandon ve Geissbuhler, 2004), sanat tarihi (Chen, 2001a, 2001b; Holt, Weiss, Niblack, Flickner ve Petkovic, 1997) gibi alanlar bu tip araştırmalara konu olan alanlardır. Başka bir çalışmada da (Gudivada ve Raghavan, 1995, s. 18) görüntü erişim sistemleri ile ilgili pratik fayda sağlayan disiplinler olarak yukarıdakiler sayılmakta ve bunlara mimarlık, mühendislik gibi alanlar eklenmektedir.

Konu ile ilgili literatür incelendiğinde 1990'lı yılların başı ile birlikte özellikle üzerinde durulan yaklaşımın içerik tabanlı görüntü erişimi olduğu görülmektedir (Huang ve Dai, 2004, s. 82). Geniş kapsamlı görsel materyal içeren koleksiyonların düzenlenmesi için metin tabanlı ve elci sistemlerin kullanımının yetersiz olacağı görülmüş ve görsel malzemelerin kendi içerikleriyle, örneğin renk, doku ya da şekilleri ile dizinlenmesi konusunda çalışmalar (Andreou ve Sgouros, 2005; Androustos, Kushki, Plataniotis ve Venetsanopoulos, 2005; Colombo ve Del Bimbo, 2002; Goodrum, Rorvig, Jeong ve Suresh, 2001) yoğunlaşmıştır.

İçerik Tabanlı Görüntü Erişim Sistemleri

Genel olarak içerik tabanlı görüntü erişim sistemleri üç temel bileşen içermektedir. Bunlar görsel öznitelik çıkarımı (*visual feature extraction*), yüksek boyutlu dizinleme (*high-dimensional indexing*) ve erişim sistemi tasarımı (*retrieval system design*) olarak adlandırılmaktadır (Rui, Huang ve Chang, 1997). Bunlar arasında görsel öznitelik çıkarımı, içerik tabanlı görüntü erişiminin temeli olarak kabul edilmektedir (Zhang ve Chen, 2002, s. 260). Öznitelik çıkarımı sayesinde görsel nesnelerin renk, doku ve şekil özniteliklerinin sisteme aktarılması ve bunların görsel içerikleri ile dizinlenmesi sağlanmaktadır (Park, Ryu, Song ve Ahn, 2005, s. 206).

Yüksek boyutlu dizinleme dizinlenecek malzemenin karmaşık yapısının bilgi erişim sistemine uyarlanması için geliştirilmiş bir dizi tekniği ifade etmektedir. Bu teknikler temelde iki noktada odaklanmaktadır. Birincisi genelde yüksek boyutlu olan öznitelik vektörlerinin boyutlarının azaltılması (*dimension reduction*) ikincisi ise insan algılamasına yönelik, görsel benzerlik hesapla-

malarını içeren çok boyutlu dizinleme (*multi-dimensional indexing*) tekniklerinin geliştirilmesidir (Rui ve diğerleri, 1999, s. 46). Boyut azaltımı, erişim kalitesi ve kolaylığını etkilemeyecek şekilde öznitelik vektörleri üstünde çeşitli matematiksel işlemler³ uygulamak suretiyle boyutlarının düşürülmesidir. Çok boyutlu dizinleme alanına aktif olarak katkı yapan üç önemli araştırma alanı bulunmaktadır. Bunlar hesaba dayalı geometri, veri tabanı yönetimi ve görüntü tanımadır (*pattern recognition*). Burada kullanılan yöntemler, dizinleme yöntemlerini benzerlik ölçümlerine (*similarity measures*) göre uyarlama yönünde kümeleme (*clustering*) ve sinir ağları (*neural nets*) gibi teknikleri içermektedir (Rui ve diğerleri, 1997).

Görüntü erişim sistemleri genellikle aşağıda yer alan özelliklerin birini ya da daha fazlasını desteklemektedir. Bu özellikler şunlardır (Rui ve diğerleri, 1999, s. 48):

- Rastgele göz gezdirme,
- Örnek kullanarak arama yapma,
- Taslak çizerek arama yapma,
- Metin ile arama yapma,
- İsteğe göre uyarlanmış görüntü kategorileri arasında dolaşma.

Görüntü erişim sistemleri temelde aynı mantık üzerinde kurulmuş olup aynı amacı gerçekleştirmek için çalışsa da, çeşitli yönlerden farklılık göstermektedir: IBM Almaden Araştırma Merkezi tarafından geliştirilen QBIC (*query by image content*) ilk ticari içerik tabanlı görüntü erişim sistemi olarak karşımıza çıkmaktadır. QBIC sistem yapısı ve kullandığı teknikler itibariyle daha sonra geliştirilen görüntü erişim sistemlerini etkilemiştir (Marques ve Furht, 2002, s. 26). Örneğin Virage Inc. tarafından geliştirilen Virage isimli içerik tabanlı görüntü erişim sistemi QBIC'e benzer Virage, QBIC tarafından yapılan işlemleri biraz daha geliştirmiş ve sorgulardaki renk, kompozisyon, doku ve yapı öğelerinin aramadaki ağırlıklarının kullanıcı tarafından ayarlanması olanağını sunmuştur (Rui ve diğerleri, 1999, s. 49). Zaman içinde pek çok farklı sistem geliştirilmiştir ve geliştirilmeye devam edilmektedir. Geliştirilen içerik tabanlı görüntü erişim sistemlerine Blobworld, Chabot, ImageRover, RetrievalWare, Photobook, VisualSEEK, Viper, Netra, MARS ve ART MUSEUM örnek olarak gösterilebilir (Marques ve Furht, 2002, ss. 25-26; Rui ve diğerleri, 1999, ss. 49-50; Şahin, 2003, ss. 10-12). Literatürde söz konusu sistemlerin değerlendirildiği ve birbirleriyle karşılaştırıldığı birçok çalışma bulunmaktadır (Deselaers, Keysers ve Ney, 2004; Li, Wang ve Wiederhold, 2000; Squire, Müller, Müller, Marchand-Maillet ve Pun, 2001).

³Söz konusu işlemler ile ilgili ayrıntılı bilgi için bkz. Bashir, Khanvilkar, Khokhar ve Schonfeld, 2004.

Görsel Nesnelerin Özellikleri ve İnsan Algısı

Çalışmanın giriş bölümünde de belirtildiği gibi, görsel bilgi üst veri ve görsel nitelikler olarak adlandırılan iki temel bileşenden oluşmaktadır. Görsel nesneden bir dizi bilgisayar işlemi sonucunda elde edilen görsel nitelikler, kullanıcıların istedikleri görüntülere erişim sağlamalarında önemli rol oynamaktadır.

Görsel nesnelerin özellikleri temel alınarak farklı yöntemler aracılığıyla, gereksinim duyulan görüntülere erişim sağlanmaya çalışılmaktadır. Piksel karşılaştırma yöntemi de görüntü erişiminde sıkça kullanılan yöntemlerden birisidir. Piksel karşılaştırma yöntemi ile ham verinin piksel değerlerine dayanan bir dizi işlem yapılarak veri tabanında yer alan görüntü ile kullanıcının aradığı görüntü arasında fark/benzerlik bulunup bulunmadığı ölçülmektedir. Söz konusu benzerliğin oranını saptamak için de hata karelerinin ortalaması (*mean square error*) hesabı kullanılmaktadır. Aşağıda hata karelerinin ortalaması (HKO) hesaplama yöntemi verilmektedir (Li, Chen ve Chen, 2003, s. 4).

$$HKO = \sum_{i=1}^x \sum_{j=1}^y \frac{|a_{ij} - b_{ij}|^2}{x \cdot y} \quad (1)$$

Denklem (1)'de geçen $x \cdot y$ çarpımı karşılaştırılan piksel bloğu içindeki toplam piksel sayısını, a_{ij} karşılaştırılan iki piksel bloğundan a-bloğu içinde yer alan pikselin değerini ve b_{ij} de b-bloğu içinde yer alan pikselin değerini ifade etmektedir.

Aşağıda, piksel karşılaştırma yöntemi kullanan bir sisteme yönlenebilecek sorgu örnekleri sıralanmaktadır (Gupta ve Jain, 1997, s. 72):

- > “Eğer portakal renginin ortalama değeri; kırmızı= 225, yeşil=130 ve mavi= 0 olarak tanımlandıysa, 100. ve 200. piksel aralıkları portakal rengi olan tüm görüntüleri bul.”
- > “Orta bölümünde, örnek olarak gösterilen resimle yaklaşık olarak aynı renge sahip olan tüm görüntüleri bul⁴.”
- > “Verilen resmin en fazla D^5 kadar kaymış hali olan bütün resimleri bul.”

Görüntü erişimde görsel nesnenin bilgi erişim sistemine aktarılabilmesi için piksel karşılaştırma yönteminin kullanımı etkin bulunmamaktadır. Çünkü kullanıcı bakış açısı ile sistemin arama ve bulma kapasitesi tam anlamıyla örtüşmemektedir. Kullanıcı gereksinimi yukarıda yer alan örneklerdeki gibi sorgular sonucunda karşılanabiliyorsa görsel bilgi için kullanılan veri modellemesi basittir. Ancak, piksel tabanlı bu modeller çeşitli sorunları da beraberlerin-

⁴Burada “orta bölüm” koordinatlarla, “yaklaşık olarak aynı renge sahip” ifadesi de renk mesafelerinin hesaplanmasıyla belirlenmektedir.

⁵D kayma mesafesini ifade etmektedir.

de getirmektedir. Öncelikle “gürültü”ye (*noise*) karşı son derece duyarlıdır. En ufak bir parazit aranan görüntülere erişimi engeller. Bunun yanı sıra taşıma (*translation*) ve döndürme (*rotation*) konusunda da bu sistemlerin eksiklikleri vardır. Talep edilen görüntüye uygun ama farklı bir açı ile görüntülenmiş bir görsel nesne sorgu sonuçları arasında yer alamaz. Renk değişimleri ve benzeri görsel durumlardaki farklılıklar, piksel değerlerini büyük ölçüde etkilediğinden yanlış sorgu sonuçlarına neden olurlar (Gupta ve Jain, 1997, s. 72).

Bu sorunlardan yola çıkarak, görsel içeriğin kullanıcı yargısına kıyasla basit fakat etkili benzerlik ölçümlerine olanak tanıyacak bir biçimde temsil edilmesi gerekmektedir. Bunun için görsel nesne özellikleri bir vektör uzayı içindeki vektörlere dönüştürülür (Fuhr, 2001).

İçerik tabanlı görüntü erişim sistemlerinin insan algısına mümkün olduğunca yakın sistemler olması için çalışılmaktadır. Bu nedenle insanların algı ya da idrak süreçlerini anlamak önemlidir. İnsanların gördükleri cisimleri algılama süreçleri içerisinde üç dürtüden söz edilebilir.

1. Genel olarak algılanan dürtü (*generally perceived stimuli*) - farkındalık yoktur,
2. Özellikle algılanan dürtü (*specifically perceived stimuli*) - farkındalık vardır,
3. Sözde-tesadüfi dürtü (*pseudo-random stimuli*) - psikolojik, sosyolojik, vb.

İnsan dürtülerinin bu türlerinden yola çıkarak görüntü erişiminde niteliksel (yüksek seviyeli) öznitelikler ve niceliksel (düşük seviyeli) öznitelikler olmak üzere iki tür öznitelik ayrımından bahsedilmektedir (Eidenberger, 2004).

Yüksek seviyeli öznitelikler, insanlar ve insanlar ya da makineler ve insanlar arasında şimdilik tam olarak çözülemeyen anlamsal uçurum (*semantic gap*) nedeniyle günümüzdeki görüntü erişim sistemlerinde fazla kullanılmamaktadır (Eidenberger, 2004). Bu uçurumun oluşmasında insanların görsel bilgi içeren bir varlığı birbirlerinden farklı algılayabilecekleri ya da tanımlayabilecekleri gerçeği etkilidir. Bunun yanı sıra, insanın kendi değişen bilgi gereksinimleri söz konusu olduğunda, farklı tanımlamalar ya da sorgu biçimleriyle sistemden aynı ya da benzer sonuçları talep etmesi durumu da bu uçurumun oluşmasında etkili olabilmektedir. Bu yüzden söz konusu sistemler düşük seviyeli özniteliklerin çıkarımında etkili olmaya çalışmakta, yüksek seviyeli özniteliklerin sisteme uyarlanması ise bu düşük seviyeli özniteliklerin güçlendirilmesi ile sağlanmaya çalışılmaktadır.

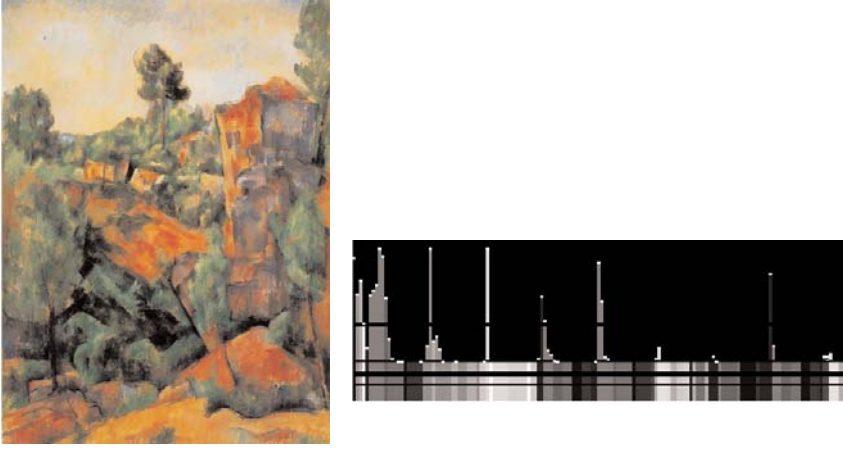
Görsel bilgi erişimde düşük seviyeli öznitelikler genel olarak renk, doku ve şekil, hareket olmak üzere üç grupta toplanmaktadır (Eidenberger, 2004; Smeulders, Worring, Santini, Gupta ve Jain, 2000). Düşük seviyeli öznitelik-

lerin üçüncüsü olan hareket ögesi bu makalenin kapsamı dışında olduğu için burada yer verilmemektedir.

Renk öznitelikleri: Ton (*hue*) ve koyuluk (*saturation*) renk ögesini niteleyen iki değişken olarak karşımıza çıkmaktadır. Ton, renklerin birlikte sergiledikleri görüntüyü, koyuluk ise ton içerisindeki saf renge ne kadar gri eklendiğini gösterir. Görüntü erişiminde renk histogramları, renk momentleri ve renk kümeleri kullanılarak benzerlik ölçümleri yapılabilmektedir (Gupta ve Jain, 1997, ss. 73-74). Şekil 1, Cézanne'a ait bir tabloyu ve söz konusu tablonun renk histogramını göstermektedir.

Şekil 1: Cézanne'a ait Tablo ve Renk Histogramı

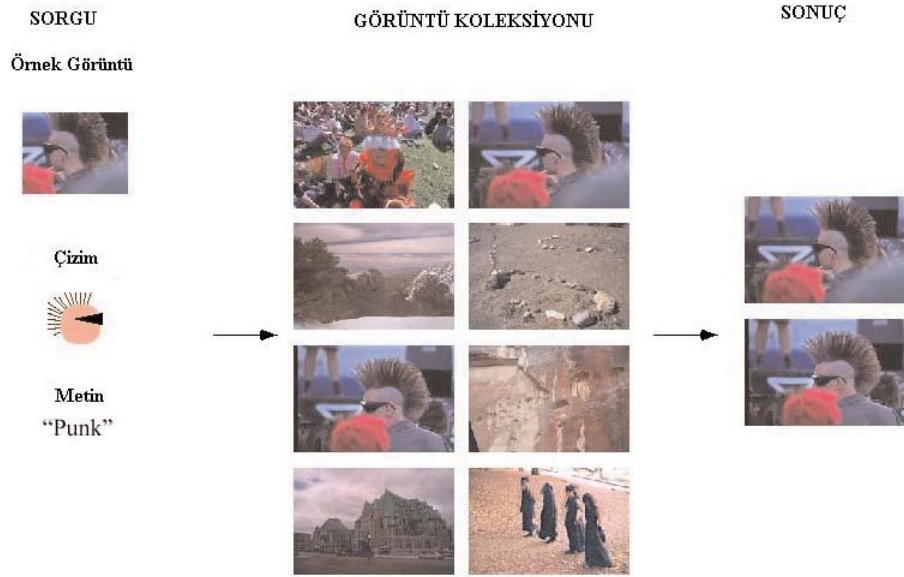
(Kaynak: Gupta ve Jain, 1997, s.70)



Doku ve şekil öznitelikleri: Doku öznitelikleri görsel nesne üzerindeki parlaklığın dağılımı ile ilgilidir ve tüm görsel yüzeylerin doğal özniteliğidir. Yüzeylerin yapısal düzeni ve çevreyle olan ilişkisi hakkında önemli bilgiler içerir. Şekil ise görsel nesnenin bölgesel ya da sınırsal alanını belirleyen yapıdır. Doku istatistiksel çerçeveyi belirlerken, şekil anlamsal çerçeveyi temsil eder. Her iki özellik de döndürme, derecelendirme gibi geometrik değişimlere karşı değişmez yapı sergilemektedir (Eidenberger, 2004).

İçerik tabanlı görüntü erişim sistemlerinde bu iki özelliğin yanı sıra metinsel bilginin de taramaya dâhil edilmesi mümkündür. Şekil 2 metin ve şekil kullanılarak yapılan bir taramayı ve bunun sonucunu göstermektedir.

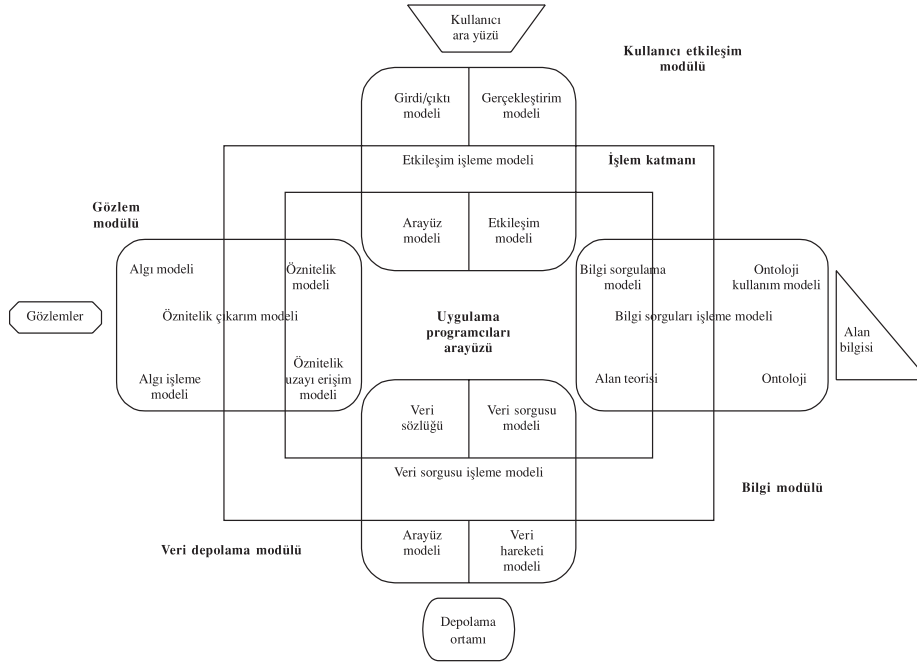
Şekil 2: İçerik Tabanlı Görüntü Erişim Sisteminde Metin ve Şekil Kullanarak Yapılan Tarama Örneği (Kaynak: Siggelkow, 2002, s. 5)



İçerik Tabanlı Görüntü Erişim Sistemlerinin Mimarisi

İçerik tabanlı görüntü erişim sistem mimarisi çok kullanıcıli geniş çaplı koleksiyonlarda verilerin ve özniteliklerin isimlendirilmiş dosyalar altında tutulduğu karmaşık bir yapı gösterir. İlgili literatür incelendiğinde Smeulders ve arkadaşlarının (2000, s. 1372) önerdiği yapı, oluşturulmak istenilen görüntü erişim sistemi için uygun görülmektedir. Söz konusu yapı, bütünlüğü olan bir sistem tasarımı için dört modülden oluşan bir iskelet önermektedir (Bkz. Şekil 3). Algısal ve özniteliksel hesaplama (*sensory and feature calculating*) modülü, alan bilgisi ve yorum (*domain knowledge and interpretation*) modülü, etkileşim ve kullanıcı arayüzü (*interaction and user interface*) modülü, depolama ve dizinleme (*storage and indexing*) modüllerinin bulunduğu bu iskeletin gelişmiş içerik tabanlı sistemlerde dengeli bir yapı oluşturulmasını sağlayacağı ifade edilmektedir. Günümüzde bu modüllerin tamamını bünyesinde bulunduran bir sisteme sahip olmanın zor olduğu bilinmekle birlikte, sistem mimarisi açısından etkinliğin artırılabilmesi için bu tip bir yapının geliştirilmesi son derece önemlidir.

Şekil 3: İçerik Tabanlı Görüntü Erişim Sistemi için Önerilen Sistem Yapısı (Kaynak: Smeulders ve diğerleri, 2000, s. 1372)













Görsel nesnelerin kendilerine has özelliklerinin birer erişim noktası olarak kullanılması, sistemin etkinliği açısından önemlidir. Doku, renk gibi öznitelikleriyle tanımlanan görsel nesneler bütünlükleri korunarak sistem içerisinde depolanmaktadır. Öznitelik vektörleri, her öznitelik vektörüne bir kayıt olacak şekilde standart doğrusal kütüklerle depolanmaktadır (Smeulders ve diğerleri 2000, s. 1369). Dizinleme ise, alan bölümlenme (*space partitioning*), veri bölümlenme (*data partitioning*), uzaklık tabanlı (*distance-based*) dizinleme teknikleri kullanılarak gerçekleştirilmektedir. Alan bölümlenme tekniğinde, özellik alanı ağaç biçimli olarak düzenlenmiştir. Bu ağaç biçimli yapıdaki her düğüm (*node*) bu alan içerisindeki bir bölgeye denk gelmektedir. Bir bölgedeki noktalar önceden belirlenmiş miktarı aştığında bölge, alt bölgelere ayrılmaktadır. Veri bölümlenme tekniği özellik alanındaki her noktayı, vektöre komşuluk eden bir bölge ile birleştirme özelliğine sahiptir. Uzaklık tabanlı dizin yapılarındaki mantık ise, bir örnek noktanın belirlenip, özellik alanının geri kalanının örnek çevresinde eşmerkezli halkalar oluşturacak biçimde bölünmesi biçimindedir (Digout ve Nascimento, 2005; Smeulders ve diğerleri, 2000, s. 1370).

Salt okunur veri tabanlarının bir adım sonrası, görsel bilgi erişim için genişletilmiş veri tabanı yönetim sistemleri olarak gösterilmektedir. QBIC ve Virage bu sistemler içinde ilk akla gelenlerdendir. Örneğin Şekil 4'te COREL görüntü veri tabanında kullanılan farklı sorgulamalara yer verilmektedir.

Şekil 4: COREL Görüntü Veri Tabanında Kullanılan Farklı Sorgu Türleri

(Kaynak: Smeulders ve diğerleri, 2000, s. 1366)

| | Örnek sorgu | Örnek sorgu sonucu |
|----------|--|--|
| tam | Uzamsal ifade / Spatial predicate  |  |
| | Görüntü ifadesi / Image predicate <i>Amount of "sky" > 20% and amount of "sand" > 30%</i> |  |
| | Grup ifadesi / Group predicate <i>Location = "Africa"</i> |  |
| yaklaşık | Uzamsal örnek / Spatial example  |  |
| | Görüntü örneği / Image example  |  |
| | Grup örneği / Group example <i>pos neg</i>  |  |

İçerik Tabanlı Görüntü Erişim Sistemlerinin Değerlendirilmesi

Geniş koleksiyona sahip görüntü erişim sistemlerinin değerlendirilmesi, metin tabanlı bilgi erişim sistemlerinde olduğu gibi gerçekleştirilmesi zor, ancak gerekli bir işlemdir. Geleneksel bilgi erişim sistemlerinde uygulanan duyarlık (*precision*) ve anma (*recall*) değerlendirme ölçütleri görüntü erişim sistemlerinde de kullanılmaktadır. Bu ölçütlerin görüntü erişim sistemlerindeki kullanımını formüle etmek için veri kümesinin D ile, sorgunun q ile, sorgu için getirilen sonuç kümesinin $A(q)$ ile ifade edildiğini düşünecek olursak: Kullanıcı özneliği açısından baktığımızda, yapılan sorguya ilişkin olarak ilgili olanların $R(q)$ ve olmayanların R şeklinde ikiye ayrıldığını söylemek mümkündür. Buradan yola çıkarak anma ve duyarlık formüllerini şu şekilde göstermek olanaklıdır (Smeulders ve diğerleri, 2000, s. 1371).

$$\text{Duyarlık} = \frac{|A(q) \cap R(q)|}{|A(q)|} \quad (2)$$

$$\text{Anma} = \frac{|A(q) \cap R(q)|}{|R(q)|} \quad (3)$$

Sözel anlatımla, sistem değerlendirilmesinde duyarlık, sorgu sonucunda elde edilen ilgili görüntülerin sayısının sorgu sonucunda gelen tüm görüntü sayısına oranıdır. Anma ise, sorgu sonucunda elde edilen ilgili görüntü sayısının sistemdeki tüm ilgili görüntülerin sayısına oranıdır [bkz. denklem (2) ve denklem (3)].

Bu ölçütler yararlı olmakla birlikte, görüntü erişimi söz konusu olduğunda, metin tabanlı bilgi erişim sistemlerinin değerlendirilmesinden daha fazla sorunla karşı karşıya kalınmaktadır. Bunun sebeplerinden birisi olarak, gelen sonuç kümesi içerisinde “ilgili” olanları seçmenin daha zor olması gösterilebilir. Burada görsel nesnenin “anlamı”nı tanımlamada karşılaşılan güçlükler ve kullanıcıların farklı bakış açıları ön plana çıkmaktadır. Dilsel içerikten bağımsız olarak yapılan yorumlama, insan algısının devreye girmesi bu zorlukların oluşmasında etkilidir.

Bir diğer sorun ise görüntü erişim sistemlerinin ilgililik ölçütüne göre değil, sadece sıralı bir liste veya sorgu sonuçları arasındaki ilişkilerin ön planda olduğu bir sonuç listesi getirmesidir. Bu da formülleri uygulayabilmek için gerekli olan değerlendirmenin yapılmasını güçleştirmektedir.

Duyarlık ve anma ölçütleri bu sorunlara rağmen birçok görüntü erişim sistemi için kullanılabilir. Bu ölçütlerin özellikle, görüntüler için etiketlerin kullanıldığı ya da metinsel tanımlamaların yapıldığı veri tabanlarında kullanışlı olduğu söylenebilir. Ancak insan algısının bu derece önemli olduğu sistemlerde performans değerlendirmesi yapabilmek için sosyal bilimlerden uyarlanacak yöntemlerle görsel bilgi erişim sisteminin değişik bölümlerinin ya da tamamının değerlendirilmesi de mümkündür (Smeulders ve diğerleri, 2000, s. 1372).

Sonuç

Görüntü erişimi için oluşturulan sistemler günümüzde metin tabanlı olmaktan çıkıp içerik tabanlı hale gelmeye başlamıştır. Kullanılan malzemenin doğasından kaynaklanan sorunlar nedeniyle tam olarak etkin bir bilgi erişim sistemi oluşturulamamaktadır. Karşı karşıya bulunulan sorunların kökeninde

makinelere tarafından tam olarak taklit edilemeyen insan algısı sonucu ortaya çıkan görsel anlayış yatmaktadır. Bu bağlamda görüntü erişimle ilgili temel sorunlar olarak farklı kullanıcı beklentileri, karmaşık kullanıcı arayüzleri, tatmin etmeyen sorgulama performansı, sisteme yöneltilen sorgulardan elde edilen düşük kalitede sonuç ve değerlendirme yöntemlerinin eksikliği karşımıza çıkmaktadır. Söz konusu problemlerin çözümü farklı disiplinlerden uzmanların bir araya gelmesiyle gerçekleştirilecek projelere dayalı oluşturulacak yeni sistemlerle mümkündür.

QBIC ve benzeri veri tabanları, veri ile sorgulama sistemi arasında dar bir veri değişim kanalı oluşturmaktadır. Var olan veri tabanına görüntü erişim modeli eklenerek oluşturulan, görselleştirme teknikleri zayıf ve bilgi yönetiminde yoksun bu sistemlerde genellikle performans düşük olmaktadır. Kullanıcılar daha gelişmiş sistemlere gereksinim duymaktadırlar. Gelecekte meydana getirilecek sistemlerin mevcut özelliklerine ek olarak kullanıcı etkileşiminin yüksek olduğu, kullanıcı etkileşiminden ortaya çıkan özelliklerle ilgili peş peşe işlemler gerçekleştirebilen, veri düzenlemesini geri plana atıp kullanıcı ya da kaynak bilgisini ön plana çıkarabilen bir yapı sergilemesi gerekmektedir.

Ülkemizde de konu ile ilgili bazı çalışmaların yapıldığı bilinmekle birlikte, pratik yarar sağlayacak ve sorunların çözümüne yol gösterecek projelere gereksinim duyulmaktadır. Görsel bilgi erişimle ilgili sorunların üzerine gidilmesi ile kullanıcıların zaman ve emek kaybı en aza indirgenecek ve görsel bilgi ile ilgilenen kişilerin daha verimli çalışabilmelerine olanak tanınacaktır.

Kaynakça

- Andreou, I. ve Sgouros, N.M. (2005). Computing, explaining and visualizing shape similarity in context-based image retrieval. *Information Processing & Management*, 41(5), 1121-1139.
- Androustos, P., Kushki, A., Plataniotis, K.N. ve Venetsanopoulos, A.N. (2005). Aggregation of color and shape features for hybrid query generation in content based visual information retrieval. *Signal Processing*, 85(2), 385-393.
- Bashir, F., Khanvilkar, S., Khokhar, A. ve Schonfeld, D. (2004) Multimedia systems: Content based indexing and retrieval. Wai Chen (ed.), *The electrical engineering handbook* içinde Boston: Elsevier Academic Press 16 Temmuz 2005 tarihinde <http://mia.ece.uic.edu/~papers/publications/EE-Handbook-Chapter-Indexing.pdf> adresinden erişildi.
- Cawkell, T. (2000). Image indexing and retrieval by content. *Information Services & Use*, 20(1), 49-58.

- Chen, H. (2001a). An analysis of image queries in the field of art history. *Journal of the American Society for Information Science and Technology*, 52(3), 260-273.
- Chen, H. (2001b). An analysis of image retrieval tasks in the field of art history. *Information Processing & Management*, 37(5), 701-720.
- Colombo, C. ve Del Bimbo, A. (2002). Visible image retrieval. V. Castelli ve L. Bergman (ed.), *Image databases: Search and Retrieval of Digital Imagery* içinde (ss. 11-33). New York: Wiley.
- Deselaers, T., Keyzers, D. ve Ney, H. (2004). Classification error rate for quantitative evaluation of content-based image retrieval systems. *ICPR*, 2(2), 505-508.
- Digout, C. ve Nascimento, M.A. (2005). High-dimensional similarity searches using a metric pseudo-grid. Proceedings of the 1st IEEE International Workshop on Managing Data for Emerging Multimedia Applications, Tokyo. 14 Temmuz 2005 tarihinde <http://www.cs.ualberta.ca/~mn/Papers/emma2005.pdf> adresinden erişildi.
- Eidenberger, H. (2004). A new perspective on visual information retrieval. SPIE Electronic Imaging Symposium, San Jose. 10 Temmuz 2005 tarihinde <http://www.ims.tuwien.ac.at/media/documents/publications/ei2004-vir.pdf> adresinden erişildi.
- Fuhr, N. (2001). Information retrieval methods for multimedia objects. *State-of-the-Art in Content-Based Image and Video Retrieval* içinde (ss.191-212). 14 Temmuz 2005 tarihinde <http://www.is.informatik.uni-duisburg.de/bib/pdf/ir/Fuhr:00b.pdf> adresinden erişildi. Boston: Kluwer.
- Goodrum, A.A. (2000). Image information retrieval: An overview of current research. *Informing Science*, 3(2), 63-67.
- Goodrum, A., Rorvig, M., Jeong, K-T. ve Suresh, C. (2001). An open source agenda for research linking text and image content features. *Journal of the American Society for Information Science and Technology*, 52(11), 948-953.
- Goodrum, A. ve Spink, A. (2001). Image searching on the Excite web search engine. *Information Processing & Management*, 37(2), 295-311.
- Gudivada, V.N. ve Raghavan, V.V. (1995). Content-based image retrieval systems. *Computer*, 28(9), 18-22.
- Gupta, A.ve Jain, R. (1997). Visual information retrieval. *Communications of the ACM*, 40(5), 70-79.

- Holt, B., Weiss, K., Niblack, W., Flickner, M. ve Petkovic, D. (1997). The QBIC Project in the Department of Art and Art History at UC Davis. Digital Collections: Implications for Users, Funders, Developers, and Maintainers, ASIS 97 içinde Washington, DC 10 Temmuz 2005 tarihinde <http://www.asis.org/annual-97/holt.htm> adresinden erişildi.
- Huang, P.W. ve Dai, S.K. (2004). Design of a two-stage content-based image retrieval system using texture similarity. *Information Processing & Management*, 40, 81-96.
- Lee, E., Whalen, T., Sakalauskas, J., Baigent, G., Bisesar, C., McCarthy, A., Reid, G. ve Wotton, C. (2004). Suspect identification by facial features. *Ergonomics*, 47(7), 719-747.
- Lehmann, T.M., Meinzer, H.P. ve Tolxdorff, T. (2004). Advances in biomedical image analysis: Past, present and future challenges. *Methods of Information in Medicine*, 43(4), 308-314.
- Lehmann, T.M., Plodowski, B., Spitzer, K., Wein, B.B., Ney, H. ve Seidl, T. (2004). Extended query refinement for content-based access to large medical image databases. *Proceedings of SPIE*, 5371, 90-98.
- Li, J., Wang, J.Z. ve Wiederhold, G. (2000). IRM: integrated region matching for image retrieval. *Proceedings of the Eighth ACM International Conference on Multimedia* içinde (ss. 147-156). New York: ACM Press.
- Li, K-F., Chen, T-S. ve Chen, K-H. (2003). Fractal image process based image comparison search engine. *Journal of Information Science and Engineering*, 19,1-14.
- Markkula, M. ve Sormunen, E. (1998). Searching for photos-journalists' practices in pictorial IR. Challenge image retrieval: A Workshop and Symposium on Image Retrieval, Electronic Workshops in Computing içinde Newcastle: The British Computer Society. 10 Temmuz 2005 tarihinde <http://ewic.bcs.org/conferences/1998/imageret/papers/paper8.pdf> adresinden erişildi.
- Markkula, M. ve Sormunen, E. (2000). End-user searching challenges indexing practices in the digital newspaper photo archive. *Information Retrieval*, 1(4), 259-285.
- Marques, O. ve Furht, B. (2002). MUSE: A content-based image search and retrieval system using relevance feedback. *Multimedia Tools and Applications*, 17, 21-50.
- Müller, H., Michoux, N., Bandon, D. ve Geissbuhler, A. (2004). A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. *International Journal of Medical Informatics*, 73(1), 1-23.

- Park, W.-B., Ryu, E.-J., Song, Y.-J. ve Ahn, J.H. (2005). Visual feature extraction under wavelet domain for image retrieval. *Key Engineering Materials*, 277-279, 206-211.
- Price, G. (2005). Yahoo increases size of image database, new features added. 24 Haziran 2005 tarihinde <http://blog.searchenginewatch.com/blog/050223-010101> adresinden erişildi.
- Rui, Y., Huang, T.S. ve Chang, S.-F. (1997). Image retrieval: Past, present and future. International Symposium on Multimedia Information Processing, Taiwan. 10 Temmuz 2005 tarihinde <http://www-db.stanford.edu/pub/gio/biblio/imagesurvey98.ps> adresinden erişildi.
- Rui, Y., Huang, T.S. ve Chang, S. (1999). Image retrieval: Current techniques, promising directions and open issues. *Journal of Visual Communication and Image Representation*, 10, 39-62.
- Siggelkow, S. (2002). Feature histograms for content-based image retrieval. PhD Thesis. Albert-Ludwigs-Universität Freiburg, Freiburg. 15 Temmuz 2005 tarihinde <http://www.doc.ic.ac.uk/~xh1/Referece/image-retrival/Feature-Histograms-for-Content-Based-Image-Retrieval.pdf> adresinden erişildi.
- Smeulders, A.W.M., Worring, M., Santini, S., Gupta, A. ve Jain, R. (2000). Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(12), 1349-1380.
- Squire, D., Müller, H., Müller, W., Marchand-Maillet, S. ve Pun, T. (2001). Design and evaluation of a content-based image retrieval system. S.M. Rahman (ed.), *Design management of multimedia information systems: Opportunities and challenges* içinde (ss 125-151). Hershey: Idea Group Publishing.
- Şahin, P. (Duygulu). (2003). *Translating images to words: A novel approach for object recognition*. PhD Thesis. Middle East Technical University, Ankara. 15 Temmuz 2005 tarihinde <http://www.cs.bilkent.edu.tr/~duygulu/papers/PhDthesis.pdf> adresinden erişildi.
- Tonta, Y. (2001). Bilgi erişim sorunu. "21. Yüzyıla Girerken Enformasyon Olgusu": Ulusal Sempozyum bildirileri, 19-21 Nisan 2001 içinde (ss. 195-206). Ankara: Türk Kütüphaneciler Derneği.
- Wu, J.K. ve Narasimhalu, A.D. (1994). Identifying faces using multiple retrievals. *IEEE Multimedia*, 1(2), 27-38.
- Zhang, C. ve Chen, T. (2002). An active learning framework for content-based information retrieval. *IEEE Transactions on Multimedia*, 4(2), 260-268.