

İmgeden Kaynak Tanılama

Emir Ertun

İsmail Avcıbaş

^{1,2}Elektrik-Elektronik Mühendisliği Bölümü, Başkent Üniversitesi, Ankara
e-posta: eertun@gmail.com e-posta: avcibas@baskent.edu.tr

Özetçe

Günümüzde sayısal kameranın cep telefonuna bütünleşmesi ile birlikte sayısal imgelerin elde edilmesi çok kolaylaşmıştır. İmgeler adli kanıt olarak kullanılmak istendiğinde ilk yanıtlanması gereken sorulardan biri kanıt imgenin gerçekten o cep telefonu ile elde edilip edilmediğidir. Bu bildiriye sayısal imgedeki bit düzlemleri arasındaki ilintinin her telefona göre farklılık göstereceği düşüncesinden yola çıkarak geliştirilen yöntem, aynı marka aynı model cep telefonlarının tanınmasında başarı ile kullanılmıştır.

1. Giriş

Klasik fotoğraf makinelerinde, adli kanıt amaçlı uygulamalarda çeşitli, kamera teşhis etme yöntemleri vardır. Bu tarz kameralar kimyasal şekilde filmle etkileşimde bulunduğundan çeşitli kamera kusurlarından yararlanarak (kusurlu lens yapısı, film iletim mekanizmasının yol açtığı negatif üzerinde saçılma izleri gibi) kaynak tespiti gerçekleştirilebilir. Günümüzde ilerleyen teknolojiyle birlikte bu tarz kameralar artık yerlerini sayısalara bırakmaktadır. Bu gelişim esnasında, sayısal imgelerin adli kanıt olarak geçerliliğini saptamakta büyük anlam ve önem kazanmıştır. Bu çalışmada amacımız cep telefonlarından JPEG formatında çekilen imgelerin hangi telefon ile çekildiğinin belirlenmesidir.

2. Sayısal İmgelerde Kaynak Tanılama Yöntemleri

2.1 JPEG Nicemleme (Quantization) ile Kaynak Tespiti

JPEG imgeler için kaynak tanılama yöntemlerinden biri nicemleme tablolarının incelenmesidir. JPEG kayıplı bir dosya sıkıştırma formatıdır. Üreticiler, modellerini kalite ve sıkıştırma ölçütlerini dengeleyecek şekilde tasarlarlar. Kurulan bu dengeler; üreticiler ve hatta aynı üreticinin ürettikleri değişik model cihazlar için farklılık gösterir. İmgeyi JPEG olarak kodlamakta kullanılan nicemleme tabloları imgeler üzerinde kendisine özgü izler bırakır. Nicemleme tablolarının incelenmesi ve karşılaştırılmasıyla imgenin kaynağının saptanmasının mümkündür [1].

2.2 Sensör Kaynaklı Gürültü Yoluyla Kaynak Tespiti

Farklı bir kaynak tespit yöntemi ise fotoalgılayıcıda bulunan fotosel yapılarının aynı şiddette ışığa verdikleri tepkilerin birbirlerinden farklı oluşudur. Bunun nedeni sensörde kullanılan silikon wafer'ın üretim hataları nedeniyle homojen olmayışıdır. Sonuç olarak aynı cihazdan elde edilen imgelerde nem ve sıcaklıktan bağımsız olacak şekilde gürültü yapısının benzerlik göstereceği öngörülmüştür [2].

2.3 Sınıflandırıcı Kullanarak Kaynak Tespiti

Bu bildiriye kullanılan yöntem ise imgelerden elde edilmiş bazı öznitelikler yardımıyla sınıflandırıcı kullanarak kaynak tespiti yapmaktır [3]. Önce kaynağı belli olan imgelerinden öğrenme yolu ile sınıflandırıcı tasarlanmakta, daha sonra eğitim aşamasında hiç kullanılmayan imgeler test edilmektedir. Sınıflandırıcıda kullanılan öznitelikler gerek zaman gerekse frekans tanım bölgesinde ikili bit düzlemleri arası ilintiyi ölçen ölçütlerdir.

2.3.1 İkili Benzerlik Ölçütleri

Kameralarda maliyeti düşürmek amacı ile, her renk kanalı için bir algılayıcı kullanmak yerine, ızgara şeklindeki renk filtresinden geçen ışık tek algılayıcının üzerine düşürülerek mozayik şeklinde imge elde edilir. Mozayik imgeden üç ayrı renk kanalının elde edilebilmesi için renk ara değerlendirme yapmak gerekmektedir. Her kamera üreticisi farklı ara değerlendirme algoritması kullandığından, oluşan imgede kameraya özgü ilinti izleri kalmaktadır. Bu çalışmada ikili doku istatistiklerinin karşılaştırılmasına dayanan ölçütler kullanıldı.

Bit düzlemlerini temsil eden vektörler $x_i = \{x_{i-k} | k = 1, \dots, K\}$ ve $y_i = \{y_{i-k} | k = 1, \dots, K\}$ olsun. Herhangi bir piksel x_i için uyuma değişkeni;

$$X_n^c = \begin{cases} 1 & \text{eger } X_c = 0 \text{ ve } X_n = 0; \\ 2 & \text{eger } X_c = 0 \text{ ve } X_n = 1; \\ 3 & \text{eger } X_c = 1 \text{ ve } X_n = 0; \\ 4 & \text{eger } X_c = 1 \text{ ve } X_n = 1; \end{cases} \quad (1)$$

$$\alpha_i^j = \sum_{k=1}^K \delta(X_i^{i-k}, j), j = 1, \dots, 4 \quad K=4 \quad (2)$$

şeklinde tanımlanabilir. Burada $\delta(m, n)$ fonksiyonu Kronecker delta fonksiyonudur:

$$\delta(m, n) = \begin{cases} 1, & m = n \\ 0, & m \neq n \end{cases} \quad (3)$$

α_i^j yi tüm imge için genellersek oluşan toplam uyuma şu şekilde hesaplanır:

$$a = \frac{1}{MN} \sum_i \alpha_i^1 \quad b = \frac{1}{MN} \sum_i \alpha_i^2$$

$$c = \frac{1}{MN} \sum_i \alpha_i^3 \quad d = \frac{1}{MN} \sum_i \alpha_i^4$$

{a,b,c,d} değişkenleri; ikili imgeler için eş-oluşum değerleri olarak yorumlanabilir. M ve N imgede bulunan yatay ve dikey piksel sayısını belirtmektedir. {a,b,c,d} değişkenleri kullanarak Tablo 1 de verilen ikili benzerlik ölçütleri tanımlanabilir [4].

Bu ölçütler kırmızı renk kanalının 7. ve 8. bit düzlemi için hesaplanmıştır.

Tablo 1: İkili benzerlik ölçütleri (I)

Bsm-1	$\frac{a}{a+b+c+d}$	Bsm-14	$\frac{bc}{ad}$
Bsm-2	$\frac{a+d}{a+b+c+d}$	Bsm-15	$\frac{ac}{bd}$
Bsm-3	$\frac{a+d}{a+2b+2c+d}$	Bsm-16	$\frac{\sqrt{(ad)-\sqrt{(bc)}}}{\sqrt{(ad)+\sqrt{(bc)}}$
Bsm-4	$\frac{a+d-b-c}{a+b+c+d}$	Bsm-17	$\frac{\sqrt{(ac)-\sqrt{(bd)}}}{\sqrt{(ac)+\sqrt{(bd)}}$
Bsm-5	$\frac{b}{a+c+d}$	Bsm-18	$\frac{(ad)-(bc)}{(ad)+(bc)}$
Bsm-6	$\frac{a}{b+c+d}$	Bsm-19	$\frac{4bc}{(a+b+c+d)^2}$
Bsm-7	$\frac{b+c}{a+d}$	Bsm-20	$\frac{a}{\sqrt{(a+b)(a+c)}}$
Bsm-8	$\frac{a+c}{b+d}$	Bsm-21	$\frac{ad}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$
Bsm-9	$\frac{a}{b+c}$	Bsm-22	$\frac{(ad)-(bc)}{\sqrt{(a+b)(a+c)(b+d)(c+d)}}$
Bsm-10	$\frac{a}{a+2(b+c)}$	Bsm-23	$\frac{a+\sqrt{ad}}{a+b+c+\sqrt{ad}}$
Bsm-11	$\frac{a}{a+0.5(b+c)}$	Bsm-24	$\frac{a}{\max((a+b),(a+c))}$
Bsm-12	$0.5\left(\frac{a}{a+b} + \frac{a}{a+c}\right)$	Bsm-25	$\frac{a}{\min((a+b),(a+c))}$
Bsm-13	$0.25\left(\frac{a}{a+b} + \frac{a}{a+c} + \frac{d}{b+d} + \frac{d}{c+d}\right)$	Bsm-26	$\frac{4(ad-bc)}{(a+d)^2 + (b+c)^2}$

Tablo 1'de verilen Bsm değerleri 7. ve 8. bit düzlemleri arasındaki farkı içerir. $bsm_i^7 - bsm_i^8$ $i=1, \dots, 26$.

Tablo 1'e ek olarak Ojala tarafından önerilmiş [5] doku ölçütleri (Tablo 2) temel alınarak oluşturulan öznelikler kullanıldı. [5]

Tablo 2: İkili benzerlik ölçütleri (II)

Ojala Kullback-Leibler Uzaklığı	$-\sum_{n=0}^{511} S_n^7 \cdot \log \frac{S_n^7}{S_n^8}$
Ojala Karşılıklı Entropi	$-\sum_{n=0}^{511} S_n^7 \cdot \log S_n^8$
Ojala Mutlak Histogram Farkı	$\sum_{n=0}^{511} S_n^7 - S_n^8 $
Ojala Minimum Histogram Farkı	$\sum_{n=0}^{511} \min(S_n^7, S_n^8)$

1	2	4
128	256 Xi	8
64	32	16

Şekil 1: Maske

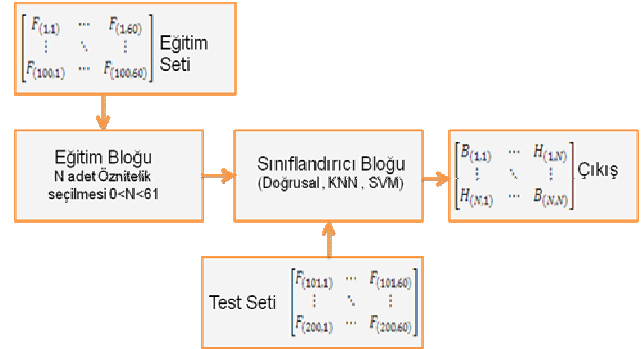
1	1	1
0	0	0
0	1	0

Şekil 2: Bit düzlemi

Şekil 1'de verilen maskeyi Şekil 2'deki bit düzlemi için hesaplayacak olursak, $S_i = 1 + 2 + 4 + 32 = 39$ olarak bulunur. Bu şekilde süzgeçleme ile elde edilen imgenin değerleri 0 ile 511 arasında değerler alacaktır. Süzgeçlenmiş imgeden hesaplanmış 512 seveli histogramlar Tablo 2'de verilen özneliklerinde hesabında kullanılmıştır.

3. Deneysel Kurgu ve Sonuçlar

Bu çalışmada iki tür deney yapılmıştır. Birinci deneyde farklı marka cep telefonları tanılanmaya, ikinci deneyde ise aynı marka ve aynı model cep telefonları tanılanmaya çalışılmıştır. Her bir cep telefonu ile 200 adet imge elde edilmiş, bu imgelerin yarısı eğitimde, diğer yarısı ise sınamada kullanılmıştır. Sınıflandırıcı olarak doğrusal, K-enyakın komşu (KNN) ve destek vektör makinesi (SVM) kullanılmıştır. Gerek uzam gerek frekans tanım bölgesinden Tablo 1 ve Tablo 2 de sıralı toplam 30'ar adet öznelik hesaplanmıştır. Özetle her bir imge 60 adet öznelikle ifade edilmektedir. Her bir cep telefonu kamerası için 200x60 öznelik matrisleri oluşturuldu. Sınıflandırma işleminde bu öznelik matrisinin 100x60'lık kısmı (ilk bölümü) eğitimde, kalan 100x60'lık bölümü test amaçlı kullanıldı.



Şekil 3: Sınıflandırma işlemi

3.1 Farklı Marka Cep Telefonlarının Üçlü Karşılaştırılması

Gerek farklı marka cep telefonları gerekse aynı marka ve model cep telefonları farklı sınıflandırıcılar ile sınıflandırılmış ve başarımları şaşırma matrisleri şeklinde Tablo 3-Tablo 14'te verilmiştir. Diyagonal üzerinde yer alan sayılar 100 imgeden kaçının doğru sınıflandırıldığını, diyagonal dışı sayılar ise yanlış sınıflandırma sayılarını vermektedir.

Bu tablolarda aynı zamanda 60 öznelikten seçilerek sınıflandırıcıda kullanılan öznelik sayıları da belirtilmektedir.

Tablo 3: Doğrusal sınıflandırıcı

Doğrusal	S.Ericsson K750i	Samsung D500	Nokia N70	% Başarım	% Genel Başarım
S.Ericsson K750i	79	11	10	79	82.6
Samsung D500	10	86	4	86	
Nokia N70	9	8	83	83	

Seçilen Öznitelikler (15)														
58	42	44	5	22	39	45	57	13	24	60	32	8	23	55

Tablo 4: KNN sınıflandırıcı

KNN N=5	S.Ericsson K750i	Samsung D500	Nokia N70	% Başarım	% Genel Başarım
S.Ericsson K750i	92	3	5	92	94.3
Samsung D500	4	94	2	94	
Nokia N70	2	1	97	97	

Seçilen Öznitelikler (6)					
53	18	45	50	37	23

Tablo 5: SVM sınıflandırıcı

SVM	S.Ericsson K750i	Samsung D500	Nokia N70	% Başarım	% Genel Başarım
S.Ericsson K750i	96	0	4	96	96.3
Samsung D500	2	95	3	95	
Nokia N70	2	0	98	98	

Seçilen Öznitelikler (5)				
38	44	13	36	37

3.2 Aynı Marka Aynı Model Cep Telefonlarının Üçlü Karşılaştırılması

Tablo 6: Doğrusal sınıflandırıcı

Doğrusal	Samsung D600 (I)	Samsung D600 (I)	Samsung D600 (I)	% Başarım	% Genel Başarım
Samsung D600 (I)	54	22	24	54	51
Samsung D600 (I)	31	55	14	55	
Samsung D600 (I)	38	18	44	44	

Seçilen Öznitelikler (5)				
60	12	30	43	51

Tablo 7: KNN sınıflandırıcı

KNN N=5	Samsung D600 (I)	Samsung D600 (I)	Samsung D600 (I)	% Başarım	% Genel Başarım
Samsung D600 (I)	61	16	23	61	57.6
Samsung D600 (I)	19	54	27	54	
Samsung D600 (I)	18	24	58	58	

Seçilen Öznitelikler (10)									
57	12	36	1	8	40	41	53	42	59

Tablo 8: SVM sınıflandırıcı

SVM	Samsung D600 (I)	Samsung D600 (I)	Samsung D600 (I)	% Başarım	% Genel Başarım
Samsung D600 (I)	69	13	18	69	64
Samsung D600 (I)	21	60	19	60	
Samsung D600 (I)	20	17	63	63	

Seçilen Öznitelikler (10)									
35	24	44	36	59	55	3	50	12	29

3.3 Cep Telefonlarının İkili Karşılaştırılması (Farklı Marka)

Tablo 9: Doğrusal sınıflandırıcı

Doğrusal	Nokia 6600	Samsung D500	% Başarım	% Genel Başarım
Nokia 6600	86	14	86	88.5
Samsung D500	9	91	91	

Seçilen Öznitelikler (12)											
12	35	23	54	31	10	17	36	21	53	51	11

Tablo 10: KNN sınıflandırıcı

KNN N=5	Nokia 6600	Samsung D500	% Başarım	% Genel Başarım
Nokia 6600	96	4	96	95.5
Samsung D500	5	95	95	

Seçilen Öznitelikler (8)							
32	60	21	36	52	18	20	5

Tablo 11: SVM sınıflandırıcı

SVM	Nokia 6600	Samsung D500	% Başarım	% Genel Başarım
Nokia 6600	99	1	99	97.5
Samsung D500	4	96	96	

Seçilen Öznitelikler (6)					
35	24	17	26	43	44

3.4 Aynı Marka Aynı Model Cep Telefonlarının İkili Karşılaştırılması

Tablo 12: Doğrusal sınıflandırıcı

Doğrusal	Samsung D600 (I)	Samsung D600 (II)	% Başarım	% Genel Başarım
Samsung D600 (I)	65	35	65	63.5
Samsung D600 (II)	38	62	62	

Seçilen Öznitelikler (16)															
38	49	5	28	1	60	34	20	10	39	47	9	42	33	26	56

Tablo 13: KNN sınıflandırıcı

KNN N=5	Samsung D600 (I)	Samsung D600 (II)	% Başarım	% Genel Başarım										
Samsung D600 (I)	76	24	76	72.5										
Samsung D600 (II)	31	69	69											
Seçilen Öznitelikler (16)														
12	33	42	41	55	5	8	16	28	37	49	40	31	26	7

Tablo 14: SVM sınıflandırıcı

SVM	Samsung D600 (I)	Samsung D600 (II)	% Başarım	% Genel Başarım										
Samsung D600 (I)	81	19	81	79.5										
Samsung D600 (II)	22	78	78											
Seçilen Öznitelikler (15)														
43	52	18	19	37	29	8	60	17	34	22	10	58	41	38

4. Vargılar

Farklı marka cep telefonları görece yüksek başarımla tanılanabilirken aynı marka ve model cep telefonlarında başarımlar düşmektedir. Başarımdaki düşüş anlaşılabilir. Destek vektör makinesi sınıflandırıcıları en yüksek başarımları vermektedir. Gelecekte aynı marka ve model tanılamada başarımların yükseltilmesine çalışılacaktır. Verilen sonuçlar sadece kırmızı renk kanalından ve sadece 7. ile 8. bit düzlemleri arasında hesaplanan öznitelikleri kullanmaktadır. Gerek diğer kanallardan, gerekse diğer bit düzlemlerinden, gerek farklı kanallar arası gerek farklı kanallar arası farklı bit düzlemlerinden hesaplanacak özniteliklerle bu başarımların artırılabilirliği düşünülmektedir.

5. Kaynakça

- [1] H. Farid “Digital Image Ballistics from JPEG Quantization” Technical Report, TR2006-583.
- [2] J. Lukas, Jessica J. Fridrich, Miroslav Goljan: Digital "bullet scratches" for images. ICIP (3) 2005: 65-68.
- [3] İ. Avcıbaşı, M. Kharrazi, N. Memon, B. Sankur, “Image Steganalysis with Binary Similarity Measures”, *Journal of Applied Signal Processing*, Vol. 17, pp. 2749-2757, 2005.
- [4] Vladimir Batagelj, Matevz Bren “Comparing Resemblance Measures”, *Distancia*, 1993.
- [5] T. Ojala, M. Pietikainen, and D. Harwood, “A comparative study of texture measures with classification based on feature distributions,” *Pattern Recognition*, vol. 29, no. 1, pp. 51–59,1996.