

Parmak Hareketlerinin Bilgisayarlı Yorumlanmasıyla Tek Oktavlı Notaların Seslendirilmesi

Vasif V. Nabiyev

*Karadeniz Teknik Üniversitesi, Bilgisayar Mühendisliği Bölümü, 61080 Trabzon, Türkiye
vasif@ktu.edu.tr*

Özet. Çalışmada çalgı aleti kullanılmadan, bilgisayarlı görü yardımıyla el durumunun (parmakların) değişimi algılanarak sanal müzik çalınması incelenmektedir. Parmak sayısının oluşturabileceği çeşitlilikler göz önünde bulundurularak sesin notaya dayalı melodisi üzerinde durulmuştur. Kamera yardımıyla alınan el görüntüsündeki parmak hareketleri bilgisayarda yorumlanarak uygun notaların seslendirilmesi sağlanmıştır. Özellik vektörü olarak el yapısının ağırlık merkezine göre parmak uçlarına olan uzaklıkların değerlendirilmesi gerçekleştirilmiştir. Böylelikle parmak hareketleri yorumlanarak piyano gibi ağır aletleri taşımadan bilgisayarda gerçek zamanlı müzik seslendirilmesinin mümkünlüğü gösterilmiştir.

Anahtar Kelimeler. İnsan-makine etkileşimi, el hareketinin tanınması, yapay zeka, müzik.

Abstract. In this paper, the task of synthesizing virtual music without adopting any musical instruments has been accomplished by detecting the changes in hand position with the help of computer vision techniques. The melody of one octave has been studied taking the possible diversities in the number of fingers into consideration. Vocalizing proper notes corresponding to the current hand position has been carried out through computerized interpretation of finger motions on a hand image recorded by a video camera. Finger positions have been determined by preprocessing the input hand image. Feature vector has been composed of the distances from hand's center of gravity to finger tips. As a result, the feasibility of real-time computerized synthesis of virtual music has been demonstrated by evaluating the finger motions without the need for heavy musical instruments such as piano.

Keywords. Human-computer interaction, hand moves recognition, artificial intelligence, music.

Received April 21, 2010; accepted December 6, 2010.

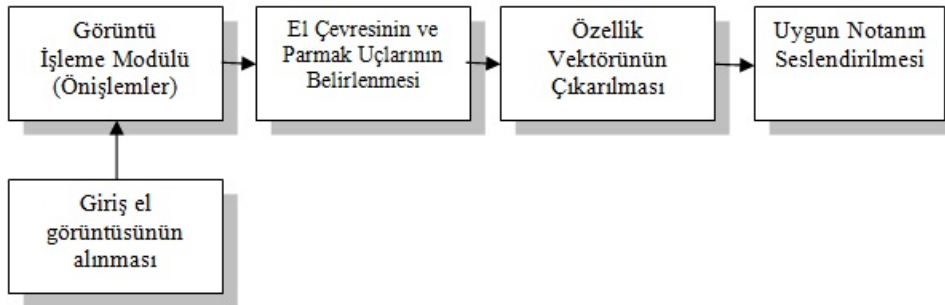
Bu makale, 24-25 Nisan 2008 tarihlerinde Çankaya Üniversitesi'nin Ankara yerleşkesinde yapılmış olan 1. Çankaya Üniversitesi Mühendislik ve Teknoloji Sempozyumu'nda sunulan ve sadece geniş bildiri özeti bölümü hakem sürecinden geçerek bu sempozyum kitapçığında yayımlanan bir makalenin revize edilmiş şekli olup Sempozyum Değerlendirme Komitesi tarafından yayımlanmak üzere Çankaya University Journal of Science and Engineering dergisine gönderilmesi önerilmiş ve derginin bağımsız hakem değerlendirmeleri sonucunda yayıma kabul edilmiştir.

1. Giriş

Güzel sanatların en soyutu olan ve kuralları gibi kullanımları da medeniyetlere göre değişen müzik, en eski sanatlardan biridir. Tuşlu müzik aletlerinden olan piyanonun sanal ortamda seslendirilmesi insan parmağının hareketleri yorumlanarak yapılabilir. Müzik notalarının tanınması, tanınan notaların seslendirilmesi [1], müzik aletlerinin seslerinin tanınması [2], müzik aleti çalanın el hareketlerinin algılanması [3], ayrıca el yapısının incelenmesi ve konumunun belirlenmesi [4,5,6,7] ile ilgili çeşitli çalışmalar yapılmaktadır.

Parmak hareketleri yorumlanarak veya virtüöz tuş takımı kullanarak piyano gibi ağır aletleri taşımadan bilgisayarda gerçek zamanlı müzik seslendirilmesi mümkün olabilir. Bu çalışmada tek oktavlı notaların seslendirilmesi için parmak hareketlerinin yorumlanması üzerinde durulmaktadır. İnsanın antropometrik yapılarından olan elin incelenmesi çeşitli araştırmalarda kullanılmaktadır. Bu araştırmalara örnek olarak *biyometri*, *kriminolojik tespit-tanıma sistemleri*, *işaret dilinin yorumlanması*, *dokunmalı ekranların tasarımı*, *animasyonlar*, *robot manipülatörleri*, *virtüöz klavyeler* gösterilebilir [8,9].

Çalışmada el hareketleri algılanarak aletsiz bilgisayarlı müzik notalarının çalınması üzerinde durulmaktadır. Kamera yardımıyla algılanan el parmağı hareketleri yorumlanarak uygun notanın seslenmesi gerçekleştirilmektedir. Çalışmanın genel yapısı Şekil 1'de gösterilmiştir. İlk önce kameradan alınan el görüntüsü Görüntü İşleme



ŞEKİL 1. Önerilen sistemin genel yapısı.

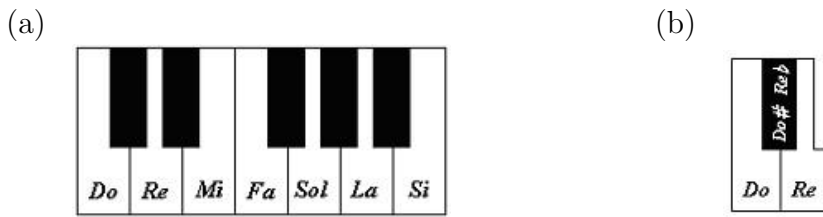
Modülüne gönderilerek önişlemler uygulanmaktadır. Sonra sınırları belirlenen gri seviye ayrık el görüntüsünden sağ ve sol el konumu, parmak uçları belirlenmektedir. Belirlenen elin ağırlık merkezinden parmak uçlarına olan uzaklıklar özellik vektörünü oluşturmaktadır. Bu özellik vektörüne göre bir oktava uygun gelen notalar seslendirilmektedir.

Sonraki bölümlerde konunun izlenmesinin kolaylığı açısından önce müzik öğeleri ve notalarına değinilmekte, sonra el görüntüsünün işlenmesi ve özelliklerinin belirlenmesi üzerinde durulmaktadır.

2. Materyal ve Metod

2.1. Müziğin öğeleri ve notalar. Müziğin başlıca bileşeni esnek cisimlerin titreşiminden doğan sestir. Sesin özellikleri, frekans, süre, genlik ve tınıyı kapsar. İnsan kulağının algılayabildiği titreşimler 16 Hz ile 20.000 Hz arasındadır. Modern piyanodaki uç notlar sırasıyla 27.5 Hz (kalın *La*) ve 4.186 Hz'e (ince *Do*) tekabül eder. Aslında çalgılardan yayılan seslerin çoğu armonik denilen, çeşitli yoğunlukta pek çok frekans içermektedir. Fakat müzik seslerinin armoniği, temel sesle sabit bir sayısal ilişki içerisinde olmaktadır [8,10].

Müzik sisteminin seslerinin yüksekliklerine göre dizilişine *sedizimi* denir. İki aynı isimli sesin belli bir mesafede bir biri ile uyum içerisinde olması ise *oktav*dır. Bilindiği üzere yedi ana sese uygun gelen *Do, Re, Mi, Fa, Sol, La, Si* tuşları sediziminde birkaç defa tekrarlanmaktadır. Aynı adlı, fakat yüksekliği farklı olan sesleri birbirinden ayırmak için, tüm sedizimi *yedi tam* ve *iki tam olmayan* oktava ayrılmaktadır. Her oktavın başlangıç sesi *Do* olarak kabul edilmektedir (Şekil 2a).

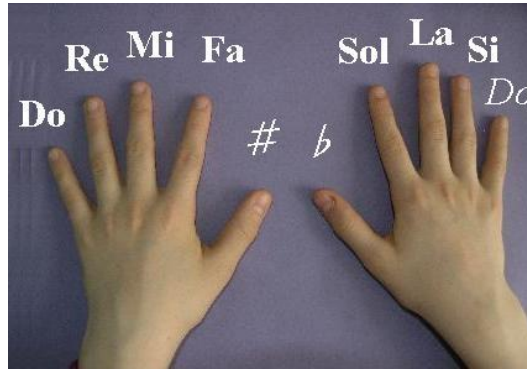


ŞEKİL 2. (a) Tam oktav, (b) Temel (*Do, Re*) ve türemiş (*Do #* ve *Re b*) tuşlar.

Piyanonun sedizimini oluşturan 88 tuştan 52'sinin beyaz, 36'sının ise siyah olduğu görülmektedir. 52 beyaz tuş yedi ana tuşun adını taşımaktadır. Diğer 36 tuş ise özel ada sahip olmamakta ve ana tuşların yarım (1/2) ton artırılması veya eksiltilmesi sonucu ortaya çıkmaktadır. Bu tuşlara *türeme* tuşlar denir. Bir oktavda 7 temel tuşun arasında 5 türemiş tuş yer almaktadır. Türemiş tuşlar komşuları olan temel notalar yardımıyla isimlerini almaktadır. Şekil 2b, temel *Do* ve *Re* arasındaki orta sesin varlığını ifade etmektedir. Bu sese uygun siyah tuş ise ismini *Do* notasının

1/2 ton artırılması (*Do #*, *Do* diyez) veya *Re* notasının 1/2 ton eksiltilmesi (*Re b*, *Re* bemol) sonucunda alır. Dolayısıyla bugünkü piyano, oktavin (*Do*'dan *Do*'ya) 12 yarım tona bölünmesinden oluşan sisteme göre akort edilir. Çalışmada 1. oktavin (orta perde) tuşlarının karşılıkları ifade edilmektedir [10].

Görüntüden elde edilen el yapısındaki parmaklar incelenerek seslendirilebilmesi için her parmağa müzik notasının uygun koyulması gerekir. Çalışmada iki elin 10 parmağına 7 temel nota ve ek olarak, diyez, bemol ve ince *Do* karşılık koyulmuştur (Şekil 3).



ŞEKİL 3. Elin parmaklarına uygun gelen temel ve türemiş notalar (*Do*, *Re*, *Mi*, *Fa*, *Sol*, *La*, *Si*, ince *Do*, # ve *b*).

Elin parmaklarının nota karşılıklarının gösterilmesi için daha kısa ifade biçimi olan harfi (*C*, *D*, *E*, *F*, *G*, *A*, *H*)[†] adlar kullanılabilir. Bunlara diyez ve bemolü ifade eden *Z* ve *B* işaretlemelerini de eklersek harfi biçimde parmak hareketlerini ifade edebiliriz. Örnek olarak aşağıda notaların harfi karşılıkları verilmiştir.

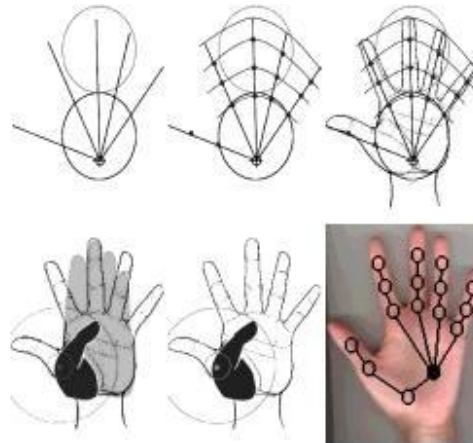
El	Parmaklar	İşaretler	Nota
Sol	Serçe parmağı	<i>C</i>	<i>Do</i>
	Yüzük parmağı	<i>D</i>	<i>Re</i>
	Orta parmak	<i>E</i>	<i>Mi</i>
	İşaret parmağı	<i>F</i>	<i>Fa</i>
Sağ	İşaret parmağı	<i>G</i>	<i>Sol</i>
	Orta parmak	<i>A</i>	<i>La</i>
	Yüzük parmağı	<i>H</i>	<i>Si</i>
	Serçe parmağı	<i>I</i>	İnce <i>Do</i>
Sol	Baş parmak	<i>Z</i>	Diyez (#)
Sağ	Baş parmak	<i>B</i>	Bemol (<i>b</i>)

Notalar	Parmak işaretleri
<i>Do Do Re Mi Re Do</i>	<i>C C D E D C</i>
<i>Do Do Do Do Re Mi Re Do</i>	<i>C C C C D E D C</i>
<i>Re Re Re Re Mi Fa Mi Re</i>	<i>D D D D E F E D</i>
<i>Re Sol Fa Mi Re Mi Re Do</i>	<i>D G F E D E D C</i>
<i>Do Do</i>	<i>C C</i>

[†]Bu harfler aslında notaların harfsel karşılıklarıdır.

2.2. El geometrisi. Parmak hareketlerinin algılanarak sanal piyanonun seslendirilmesinde el görüntüsünden parmakların konumlarının tanınması ve notalarla ilişkilendirilmesi gerekmektedir. El yapısının anlaşılması için parmakların, bileğin yerinin, elin genişliğinin, parmağın konumunun belirlenmesi önem taşır.

Elin yapısı incelendiğinde, elin yaklaşık iki eşit dikdörtgen (veya kare) alanı kapladığı görülmektedir. Bu alanlara ovaler çizilerek, alt kısımda seçilen noktadan (ilk üçü 15° olmak koşulu ile) 4 doğru geçer. 4 parmağı (işaret, orta, yüzük ve serçe) ifade eden doğrular tepesi orta parmakta olan 3 kavisle kesilir. Parmak ucuna yakın olan eklem, parmağın $1/3$ 'nin üstünde yerleşir. Geriye kalan kısım ise yarıya bölünür ve sonuçta parmağın 2 eklemi bulunur. 3. eklem ise parmakların avuçla kesiştiği kısımlarda saklıdır. Başparmak diğer parmalardan daha farklıdır ve kendi oku üzerinde hareket etmektedir. Baş parmak avuç içerisine taban noktasına eğildiğinde serçe parmağın başlangıç noktasına dokunmakta ve işaret parmağının 2. eklemi kadar dönebilmektedir. İşaret ve yüzük parmakları yaklaşık aynı uzunlukta olmakta, yüzük parmağı yaklaşık $1/2$ parmak yastığı kadar daha uzun olmaktadır. Serçe parmağı yüzük parmağının üst eklemine kadar uzanmaktadır. Serçe parmağı yaklaşık bilek eni uzunluğundadır. Baş parmak dışında geriye kalan parmakların boylarının büküm kısımlarına olan oranı yaklaşık altın orana, yani 1.62 değerine eşit olmaktadır. Ayrıca elin yapısının kare değil, daha çok kama biçiminde olduğunun bilinmesi önemlidir (Şekil 4).



ŞEKİL 4. El geometrisi.

2.3. El görüntüsünün işlenmesi. Önışlemler, alınan el görüntüsündeki gürültülerin bastırılması, kenarların algılanması, resmin ikili hale dönüştürülmesi, boşlukların doldurulması, çizgilerin inceltilmesi, uç noktaların belirlenmesi şeklinde sıralanabilir. Gürültülerin giderilmesinde aritmetik ortalamaya ve orta değere dayalı klasik filtrelemeler kullanılmaktadır. El sınırlarının belirlenmesi ve ikili durumun elde edilmesi için gri seviyeye geçilmesi gerekmektedir. Alınan el görüntülerinde her piksel kırmızı, yeşil ve maviyi (RGB) ifade eden 3 byte ile tutulmaktadır. Renklerin baskınlıkları göz önünde bulundurularak bu 3 değeri toplayıp üçe bölmek yerine toplamları 1 olan farklı katsayılarıyla çarpıp toplamak daha doğru sonuç vermektedir. Çalışmada RGB renk uzayından gri seviyeye dönüşümde mavi 0.114, yeşil 0.587 ve kırmızı 0.299 katsayılarıyla çarpılıp toplanmaktadır [11]. Katsayılarının bu şekilde belirlenmesi; insan gözünün en çok yeşil renk bileşenine, daha sonra kırmızı ve en az mavi bileşenlere duyarlı olmasından kaynaklanmaktadır. Elde edilen gri resimlerin siyah-beyaz hale getirilmesi belli bir değerin altında kalan piksellerin '0' ve üstünde kalan değerlerin '1' yapılmasıyla gerçekleştirilir. Bu değerin belirlenmesi için resmin histogram dağılımından yola çıkılır. Histogram dağılımındaki piksellerin yığıldığı iki ayrı bölgeyi ayıran orta nokta eşik değeri olarak seçilmektedir.

2.3.1. Görüntü işleme modülü. Görüntü işleme modülünde, ele ilişkin giriş görüntüsü ilk olarak 8 bitlik gri yapıya dönüştürülür. Daha sonra ikili görüntüyü elde etmek için, oluşturulan gri görüntünün histogramı çıkarılır. Histogram dağılımından hesaplanan eşik değeri yardımıyla gri görüntü ikili görüntüye dönüştürülür (Şekil 5).

Çalışmada eşik değerinin seçiminde Otsu algoritması kullanılmıştır [9,12]. Otsu algoritması, görüntüdeki piksel değerlerinin dağılımlarına göre bu piksellerin kümelendirilmesini sağlamaktadır. Görüntünün piksel değerleri $L([1, 2, \dots, L])$ gri seviye ile temsil edilirse, i . seviyedeki piksel sayısı n_i olduğunda toplam piksel sayısı aşağıdaki şekildedir:

$$N = n_1 + n_2 + n_3 + \dots + n_i + \dots + n_L.$$

İki seviyeli eşik değeri seçiminde pikseller $[1, 2, \dots, t]$ ve $[t + 1, \dots, L]$ gri seviye değerlerini içeren C_1 ve C_2 gibi iki sınıfa ayrılmaktadır.

$$p_i = n_i/N, \quad \omega_1(t) = \sum_{i=1}^t p_i \quad \text{ve} \quad \omega_2(t) = \sum_{i=t+1}^L p_i$$

olarak tanımlanırsa C_i kümeleri aşağıdaki bağıntılarla verilebilir:

$$C_1 : p_1/\omega_1(t), p_2/\omega_1(t), \dots, p_t/\omega_1(t),$$

$$C_2 : p_{t+1}/\omega_2(t), p_{t+2}/\omega_2(t), \dots, p_L/\omega_2(t).$$

Otsu metodu, sınıflar arası dağılımı (σ_B^2) maksimum yapacak optimum t^* eşik değerini belirlemektedir:

$$\sigma_B^2 = \omega_1\omega_2(\mu_2 - \mu_1)^2,$$

$$\mu_1 = \sum_{i=1}^t ip_i/\omega_1(t), \quad \mu_2 = \sum_{i=t+1}^L ip_i/\omega_2(t),$$

$$t^* = \text{ArgMax}_{1 \leq t \leq L} \{\sigma_B^2(t)\}.$$

Şekil 5'te parmak alfabesine uygun giriş görüntüsünün akıllı eşikleme aşamaları gösterilmiştir. Bulunan eşik değeri $t^*=93$ 'tür.



ŞEKİL 5. İkili görüntünün elde edilmesi aşamaları.

Bir sonraki aşamada oluşturulan ikili görüntü bir maske gibi kullanılarak, gri görüntünün arka planı elimine edilir. Bu işlemler sonucunda klasik kenar algılama yöntemlerine gerek duyulmadan el sınırları belirlenmektedir (Şekil 6).



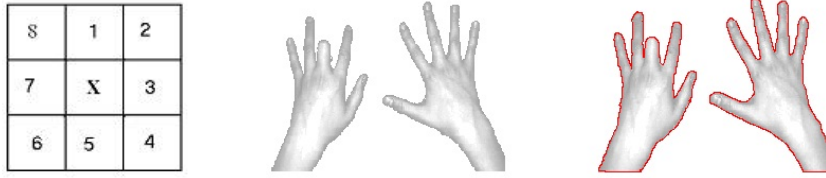
ŞEKİL 6. Ayrık el görüntüsünün elde edilmesi.

Arka plandan arındırılmış el görüntülerinin sınırlarının belirlenmesi için kontur izleme algoritması kullanılmaktadır. Bu algoritma ile ikili görüntü üzerinde başlangıç noktasından sınır izlenerek tekrar aynı noktaya geri dönlülmektedir. Nesnenin

çevresinin belirlenmesi için sekiz yönlü arama yapılmaktadır. Bu algorithmada izlenen adımlar aşağıdaki verilmiştir:

1. Yeni piksel değeri başlangıç pikseli değilken arama yapılır.
2. X pikseli için 1 yönünden aramaya başlanır. 1 yönünün seçilebilmesi için; 1'in siyah piksel, 1 için bir önceki yön sayılan 8'in beyaz piksel olması ve bir adım önce ilerleme yaptığımız yön 1 için zıt yön sayılan 5 yönü değilse, yeni konum 1 nolu piksel olur ve işlemlere bu pikselden devam edilir.
3. Eğer 2. adım sağlanmazsa bir sonraki değerli piksele geçilir.

Bu algoritma geliştirilirse; yeni piksel değeri bulunabilmesi için artan sıradaki yönlerde arama yapılır. Yön değerinin yeni değer olarak atanması için; kendi değerinin siyah, ondan bir önceki arama yönündeki pikselin beyaz olması ve bir önceki arama yönünün ona zıt olan yönlerden biri olmaması gerekir. Zıt yön değerine bakılmasının nedeni; tekrar eski değere dönmesini önlemek içindir. Şekil 7'de el çevresinin işaretlenmesinin sonuçları gösterilmiştir.



ŞEKİL 7. Sekiz yönlü arama ve uygulama sonucu.

2.3.2. Parmakların yorumlanabilmesi için özellik vektörünün belirlenmesi.

Parmak araları belirlenirken iki parmak arasında kalan siyah bölgede en alttaki noktalar seçilmektedir. Dolayısıyla 4 ara noktanın belirlenmesi gerekmektedir. Resim, üstten başlayarak satırlar boyunca taramır ve siyah aralıklar belirlendikten sonra aralıkların orta noktalarına göre parmakların avuçla kesişme noktaları seçilir. İlk ve son nokta için bölgedeki iç kısımlardan aşağıya doğru ilerlenir.

Bileğin bulunmasında satırlar boyunca aşağıdan yukarıya doğru bileğin bulunduğu muhtemel beyaz aralık belirlenir. Soldan ve sağdan ikili geçişlere göre ilk referans noktaları seçilir. Sabit adımlarla (2 piksellik) bir üst seviyedeki bilgilerle var olan bilgiler karşılaştırılır ve tekrarlı biçimde bir artışa rastlanırsa işlemler durdurularak bilek yeri tespit edilir.

Çalışmada tek elle sanal seslendirme gerçekleştirilebileceğinden elin sağ veya sol olmasının da tespiti gerekmektedir. Eğer 1. parmak arası 4. parmak arasından satır

olarak daha aşağıda ise bu sol elin işlendiğini göstermektedir. Bu yüzden sol sınır 1. parmak arasının X koordinatıdır. Sağ sınır ise 4. parmak arasının 3. parmak arasına yatay uzaklığının, 4. parmak arasının X koordinatına eklenmesi ile bulunmaktadır. Sağ el için bu işlemlerin tersi yapılmaktadır.

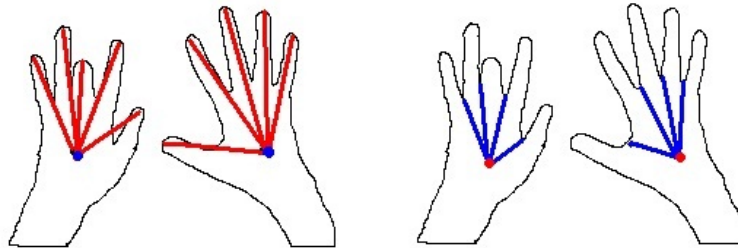
Parmakların uç noktalarını bulmak için kontur izleme algoritması ile elde edilen eli çevreleyen pikseller dizisinden faydalanılabilir. El sınırı belirlendikten sonra bulunan bu kapalı yapının ağırlık merkezi kolaylıkla bulunabilir:

$$G_x = \frac{\sum_{j=1}^{\text{en}} \sum_{i=1}^{\text{boy}} F_x(i, j)}{\text{siyah piksel sayısı}}, \quad G_y = \frac{\sum_{i=1}^{\text{boy}} \sum_{j=1}^{\text{en}} F_y(i, j)}{\text{siyah piksel sayısı}}.$$

Özellik vektörünün oluşturulmasında kontur izleme algoritması ile elde edilen dizinin elemanlarının ağırlık merkezine olan uzaklıkları hesaplanır. Bu diziyeye tepe tırmanma (hill climbing) arama algoritması uygulanarak parmakların yeri ve kapalı veya açık olması yorumlanabilir. Bu algoritma aşağıdaki şekilde özetlenebilir:

1. Başlangıç konum seçilir ve geçerli konum olarak atanır.
2. Bir sonraki konum komşu konum olarak alınır.
3. Komşu konumun değeri geçerli konumun değerinden büyükse, komşu konum geçerli konum olarak atanır.
4. Eğer geçerli konum komşu konumdan eşik değeri kadar büyükse geçerli konum yerel maksimum olarak belirlenir.

Söz konusu algoritma ele uygulandığında parmak uçlarının bulunmasında kolaylıklar sağlamaktadır. Çünkü parmak sınırlarının ağırlık merkezine göre uzaklıkları doğrusal olarak artıp azalmaktadır. Bu sınırları içeren diziler grafiksel olarak gösterilirse elde edilen pikler parmak uçlarını gösterecektir. Pikler arasındaki minimum noktalar ise parmak aralarına karşılık gelecektir (Şekil 8).



ŞEKİL 8. Tepe tırmanma algoritmasının uygulanması.

Parmak uçları maksimum, parmak araları ise ardışık şekilde minimumlara uygun gelmektedir. Bir pikselin parmak ucu olup olmadığına bakmak için pik noktalarının 15 piksel önceki ve 15 piksel sonraki değeri arasındaki uzaklık hesaplanır. Uzaklık değeri bir parmak kalınlığı şeklinde yorumlanabiliyorsa parmak ucu olarak atanır. Aşağıda uygun kod kısmı verilmiştir:

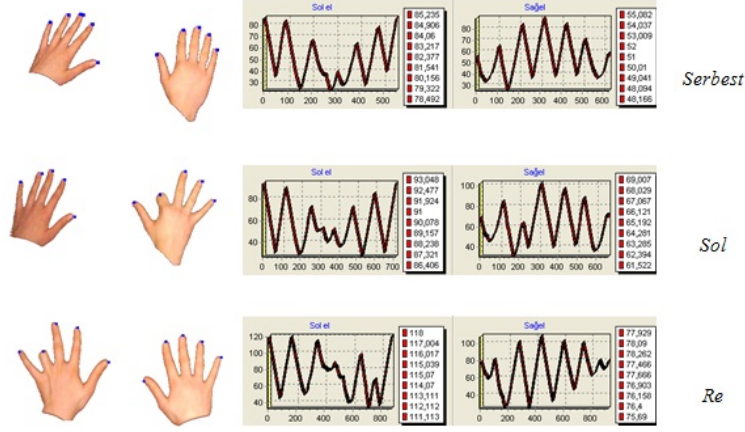
```
//iki nokta arasındaki uzaklık
double MAXY_uzaklık(int y, int x, int mxy, int mxy_x)
{
    int tempx, tempy;
    tempx = abs (mxy_x - x);
    tempy = abs (mxy - y);
    return sqrt((tempx*tempx) + (tempy*tempy));
}
.....
//indis - sınır noktasının indisi
kalınlık = MAXY_uzaklık (B[indis-15][0], B[indis-15][1], B[indis+15][0], B[indis+15][1]);
if (kalınlık < 20)
{
    // parmak ucu olarak atama yapılır.
}
```

Parmak araları da benzer bir şekilde ters mantık ile bulunur. Ancak parmak arası olarak yanlış yorumlamaların önlenmesi için ilk parmak ve son parmak ucu arasında olmayan değerler değerlendirilmez. Parmağın tuşa uygun olarak bükük olup olmadığını yorumlamak için bulunan parmak ucu ve parmak aralarının uzaklık değerleri göze alınır. Hangi parmağın bükük olduğu yorumlanarak kullanıcının hangi notayı çaldığı belirlenir. Belirlenen bükülü parmağa göre notanın seslendirilmesi gerçekleştirilir.

Şekil 9'da elin *normal* konumu, *Sol* ve *Re* notalarının karşılığı bükük sağ işaret ve sol yüzük parmaklarına uygun grafikler gösterilmiştir. Grafiklerdeki 9 değer seçilen el bölgesinin ağırlık merkezinden uygun olarak parmak uçlarına ve aralarına olan mesafeleri göstermektedir. Maksimumlar parmak uçlarına, minimumlar ise parmak aralarına uygun gelmektedir.

3. Tartışma ve Sonuçlar

Bu çalışmada el geometrisi ve parmakların konumu değerlendirilerek müzik notalarının sanal seslendirilmesi üzerinde durulmuştur. Çalışmada basitlik açısından



ŞEKİL 9. Özellik vektörüne göre parmakların yorumlanarak uygun notanın seslendirilmesi.

sesin yüksekliği ve seslenme süresi üzerinde durulmamış, genellikle notaya dayalı konumsal yapısı ele alınmıştır. Kamera yardımıyla alınan el görüntüsünden parmaklar ve bu parmakların konumu dikkate alınarak uygun nota seslendirilmesi gerçekleştirilmiştir. Uygulamalarda elin konumu ve yönü sabitlendirilmediğinde, ağırlık merkezinden yola çıkıldığından bazı hataların olduğu görülmüştür. Ayrıca bitişik parmaklar olan durumlarda parmakların konumlarının belirlenmesinde zorluklar oluşmakta ve bundan dolayı doğru notanın seslendirilmesinde hatalar ortaya çıkmaktadır. Bu hatalar elin yapısı dikkate alınarak giderilebilir. Ancak çalışmanın doğal şekilde gerçekleştirilmesi için seslerin 1 veya 1/2 ton tizleştirilmesini veya pesleştirilmesini gösteren *alterasiya* işaretlerinin de dikkate alınması gerekir. Ayrıca notalarda tüm eser boyu alternasiya işaretlerinin etkili olduğunu gösteren anahtar işareti de kullanılmaktadır. Bundan başka *inarmonik* seslerin, *diotonik* ve *kromatik* tonların dikkate alınması gerekir [10]. Örneğin piyanoda gerçek müzik için aşağı, orta ve yukarı ses perdelerine uygun gelen tüm 9 oktavın da seslendirilmesi daha gerçekçi olacaktır. Bu ise yalnız parmak değil, diğer el hareketlerinin de gelecek çalışmalarda dikkate alınması gerektiğini gösteriyor.

Kaynaklar

- [1] F. Rossant, A global method for music symbol recognition in typeset music sheets, *Pattern Recognition Letters* **23** (2002), 1129–1141.
- [2] C. Nolker and H. Ritter, Visual recognition of continuous hand postures, *IEEE Transactions on Neural Networks* **13** (2002), 983–994.

- [3] D. O. Gorodnichy and A. Yogeswaran, Detection and tracking of pianist hands and fingers, *Proceedings of the Third Canadian Conference on Computer & Robot Vision (CRV'06)* Québec, Canada, June 7-9, 2006.
- [4] Y. Bulatov, S. Jambawalikar, P. Kumar and Saurabh Sethia, Hand recognition using geometric classifiers, *Proceedings of the First International Conference on Biometric Authentication (ICBA 2004)*, Hong Kong, China, July 15-17, 753–759.
- [5] Y. Wu and T. S. Huang, Hand modeling, analysis, and recognition: For vision-based human computer interaction, *IEEE Signal Processing Magazine* **18** (2001), 51–60.
- [6] Ra L. N. Wong and P. Shi, Peg-free hand geometry recognition using hierarchical geometry and shape matching, *Proceedings of the IAPR Workshop on Machine Vision Applications* (2002) 281–284.
- [7] E. Yörük, E. Konukoğlu, B. Sankur, and J. Darbon, Shape-based hand recognition, *IEEE Transactions on Image Processing* **15** (2006), 1803–1815.
- [8] V. V. Nabiyev, *Yapay Zeka: Problemler, Yöntemler, Algoritmalar*, Seçkin Yayıncılık 2005.
- [9] V. V. Nabiyev, C. Kose, and S. Bayrak, An artificial neural network approach for sign language vowels recognition, *The International Scientific Conference of Problems of Cybernetics and Informatics*, Baku, Azerbaijan, 24-26 October 2006.
- [10] M'lou Dietzer, *First Impressions: An Intermediate Piano Method*, Alfred Music Publishing 2007.
- [11] W. K. Pratt, *Digital Image Processing*, John Wiley & Sons, New York 1991.
- [12] N. Otsu, A threshold selection method from gray-level histograms, *IEEE Transactions on Systems, Man and Cybernetics* **9** (1979), 62–66.