

**İSTANBUL TEKNİK ÜNİVERSİTESİ
ELEKTRİK-ELEKTRONİK FAKÜLTESİ**

A TOUCH GESTURE RECOGNITION LIBRARY

Bitirme Ödevi

**Soydan Ünal
040030056**

**Bölüm: Bilgisayar Mühendisliği
Anabilim Dalı: Bilgisayar Bilimleri**

Danışman: Yrd. Doç. Dr. Feza Buzluca

Mayıs 2008

**İSTANBUL TEKNİK ÜNİVERSİTESİ
ELEKTRİK-ELEKTRONİK FAKÜLTESİ**

A TOUCH GESTURE RECOGNITION LIBRARY

Bitirme Ödevi

**Soydan Ünal
040030056**

**Bölüm: Bilgisayar Mühendisliği
Anabilim Dalı: Bilgisayar Bilimleri**

Danışman: Yrd. Doç. Dr. Feza Buzluca

Mayıs 2008

Özgünlük Bildirisi

1. Bu çalışmada, başka kaynaklardan yapılan tüm alıntıların, ilgili kaynaklar referans gösterilerek açıkça belirtildiğini,
2. Alıntılar dışındaki bölümlerin, özellikle projenin ana konusunu oluşturan teorik çalışmaların ve yazılım/donanımın benim tarafımdan yapıldığını bildiririm.

İstanbul, 20 Mayıs 2008

Soydan Ünal

İmza:

A TOUCH GESTURE RECOGNITION LIBRARY

(SUMMARY)

By the definition; A **Touchscreen** is a display which can detect the location of touches within its display area. Interpretation of these touches for user interface navigation offers a much more intuitive user experience than what was available with traditional hardware such as mouse and keyboard.

Especially in the last few years, devices with touchscreen have become more accessible to end users due to advancements in production technology and fall in prices. Also; their capabilities have improved vastly, most of today's devices offer a pixel-perfect accuracy and some even can detect multi-touch. Examples of popular devices with a touchscreen could be the iphone(cell-phone), Nintendo DS.(hand-held game console), Newton(PDA), ipod Touch(mp3 player), Microsoft Surface(Table-surface computer).

The aim of this project was developing a complete library for recognition of complex shapes and gesture input on a touch-screen. Most devices already have gesture recognition capabilities for simple gestures such as a click, drag, tap, long press etc. Enhancing these capabilities and providing recognition for shapes and characters would allow an even better user experience. Although there are several commercial and professional projects on this topic; most of them aim for handwriting recognition and handwritten text input rather than user interface navigation. That being said; recognition capabilities of this project are also not trivial and it can be used for handwriting recognition.

As devices with limited screen size would benefit the most from a gesture-based navigation system, this project was developed mainly for mobile devices. However; it can also be used on other platforms, such as for detecting mouse gestures on a PC. It was implemented on the newly announced Google Android Platform which has an open and free API, thus this project can also be considered as an exploration of said platform.

This project has also been submitted to Android Developer Contest round I under the name "Dkn Touch Gesture Recognition Library". Presentation videos can be seen below:

User interface navigation: http://www.youtube.com/watch?v=JxwGMtNS_YA

Handwriting recognition: <http://www.youtube.com/watch?v=C7A6OHwfSOE>

A TOUCH GESTURE RECOGNITION LIBRARY

(ÖZET)

Projenin Türkçe adı “Dokunmatik ekranda şekil ve işaret tanınması üzerine bir kütüphane” olarak düşünülmüştür.

Tanım itibari ile dokunmatik ekran; görüntü alanı üzerindeki dokunuşları algılayabilen bir donanımdır. Bu dokunuşlarla ilgili konum, basınç gibi bilgiler ekranın bağlı bulunduğu aygıt tarafından yorumlanarak bir kullanıcı giriş/çıkış metodu sağlanır. Bu tip bir kullanıcı arayüzünün, şu an yaygın olarak kullanılan klavye ve fare gibi donanımlara nazaran daha doğal bir kullanıcı deneyimi ortaya koyduğu söylenebilir.

Özellikle son yıllarda üretim teknolojilerindeki gelişim eşliğinde dokunmatik ekranların maliyeti düşmüş ve dolayısıyla dokunmatik ekrana sahip cihazların piyasadaki yaygınlığı artmıştır. Aynı zamanda ekranların kabiliyetleri de muazzam düzeyde yükselmiştir, şu an piyasadaki birçok cihaz piksel hassasiyetinde dokunma konumu tespiti yapabilmekte, hatta bir kısmı aynı anda birden fazla dokunuşu tespit edebilmektedir. Günümüzün popüler dokunmatik ekranlı cihazlarına örnek olarak iphone(cep telefonu), Nintendo DS(taşınabilir oyun konsolu), Newton(PDA), ipod Touch(mp3 çalar) ve Microsoft Surface (henüz prototip aşamasında bir masa yüzeyi bilgisayarı) gösterilebilir.

Bu projenin amacı; dokunmatik ekranlar üzerinde karmaşık hareketlerin tanımlanmasına yarayan bir yöntem ve kütüphane ortaya koymaktır. Dokunmatik ekrana sahip aygıtların çoğu zaten tek basış, art arda basış, basılı tutma, sürüklenme benzeri temel hareketleri tanımlama yeteneğine sahiptir. Projenin amacı ise; “karmaşık hareket” olarak niteleyebileceğimiz yazım karakterleri(harf, rakam) gibi görsel şekilleri tanıyan bir sistem ortaya koymaktır.

Bu tip karmaşık hareketlerin tanınmasını gerçekleştiren profesyonel ticari yazılımlar piyasada mevcut olsa da; bunların büyük kısmı el yazısı tanıma üzerinedir ve daha çok özel yazım kalemine sahip PDA cihazlarını hedeflemektedir. Bu yazılımların atası olarak; taşınabilir bir cihaz üzerinde ilk kez bu tip bir işlevi gerçekleştiren Palm firmasının Graffiti yazılımı gösterilebilir. Takip eden dönemlerde bu konu üzerinde yeni yazılımlar ortaya konmuş ve el yazısı tanıma işlevine sahip not tableti şeklinde yeni cihazlar (Apple Newton Messagepad gibi) piyasaya çıkmıştır. Fakat bu cihazlar beklenen düzeyde ilgi görmemiş ve fazla yaygınlaşamamıştır. Günümüz itibari ile Palm firmasının PDA cihazlar arasındaki pazar payı giderek düşmekte olsa da Graffiti ve benzeri tanıma yazılımları Windows Mobile, Symbian OS gibi platformlarda varlığını sürdürmektedir.

Bu projede; el yazısı tanımadan ziyade parmak ile çizilebilecek temel şekillerin tanınması ve bu şekilleri temel alan bir kullanıcı arayüzü modelinin ortaya koyulması hedeflenmiştir. Proje özel yazım kalemine sahip PDA’lardan ziyade parmak ile kullanılan dokunmatik ekrana sahip yeni nesil mobil cihazlar için düşünülmüştür. Ortaya konacak kullanıcı arayüzü; hareketleri işlemlere atamak ve bu sayede kısayollar oluşturmak sureti ile çalışacaktır. Buna örnek olarak;

ekrana “t” simgesi çizildiğinde internet tarayıcının açılması verilebilir. Böyle bir arayüz modelinin; butonların ekranda görece olarak büyük bir alanı bloke ettiği küçük ekranlı mobil cihazlarda kullanışlı olacağı düşünülmektedir. Öte yandan; ortaya konulan projenin şekil tanıma yeteneği el yazısı için de kullanılabilir düzeyde yüksektir. İlerleyen kısımlarda hem arayüz kullanımı hem de el yazısı tanıma denemelerinin izlenebileceği video bağlantıları verilecektir.

Proje; yenilikçi bir mobil platform olarak ortaya konan Google Android işletim sisteminde geliştirilmiştir. Bu platformun seçiminde hem açık ve bedava bir API’ye sahip olması; hem de sistem uygulaması – sonradan yüklenen ayrımı yapmadan tüm uygulamalara (belli bir güvenlik modeli dahilinde) eşit erişim olanakları vermesi etken olmuştur. Bu sayede; sadece tek bir program içinde değil; işletim sistemi düzeyinde hareketleri kısayol olarak kullanmak mümkün olacaktır.

Android Platformu henüz prototip aşamasında olduğu için; proje bu platformun özelliklerinin denemeye yönelik bir girişim olarak da değerlendirilebilir. Android platformuna sahip gerçek cihazlar piyasaya henüz çıkmadığından, geliştirme aşamasında prototip SDK ve emülatör kullanılmıştır. Bu nedenle; emülatördeki başarımlar üzerinden gerçek bir dokunmatik ekrandaki başarımının kestirilemeyeceği düşünülebilir. Öte yandan; projenin tanıma ile ilgili modülleri platform ve giriş-çıkış aygıtından zaten bağımsız olarak tasarlanmıştır. Bu modüller sadece 2 boyutlu düzlem üzerinde zamanla değişen x-y koordinatlarını giriş olarak kabul etmekte ve sonuçları bu bilgiye göre üretmektedir. Hal böyleyken; dokunmatik ekran, fare ya da herhangi diğer konum bilgisi dönebilen giriş çıkış aygıtı (örneğin bir sensör) ile hareket tanıma mümkün olmaktadır. Bu sayede projenin herhangi diğer bir platforma çevrilmesi de kolaylaşmaktadır.

Toplayacak olursak; projeye başlarken ortaya konan hedefler şu başlıklar altında özetlenebilir:

- Şekil ve hareket tanımayı kabul edilebilir başarımlar düzeyinde gerçekleştiren bir kütüphanenin ortaya konması; bu hedef doğrultusunda şekil tanıma üzerine bilimsel makalelerin gözden geçirilerek uygun bir algoritmanın kodlanıp hayata geçirilmesi.
- Şekil tanıma işlemlerinin; sınırlı işlemci ve hafıza kaynaklarına sahip bir mobil cihazda gerçekleşeceği göz önüne alınarak optimize hale getirilmesi. Hız ve tanıma başarımı arasında uygun bir denge düzeyinin tespit edilmesi.
- Şekil tanımanın kullanıcı arayüzleri ile birleştirilmesi ve ortaya bir model konulması. Bu modelin denenmesi, örnek uygulamalar gerçekleyerek kullanılabilirliğinin araştırılması.
- Çeşitli işlevler için uygun düşen ve doğal bir kullanıcı deneyimi sağlayan şekil ve hareketlerin önerilmesi. (Örneğin saat yönünde çizilen bir çemberin sesi açması, ters yönde çizilen çemberin ise kapaması)

Ayrıca proje; “Dkn Touch Gesture Recognition Library” adıyla, Android Platformunun tanıtımı için gerçekleştirilen “Android Developer Contest” yarışmasına da katılmıştır. (Dkn ismi “dokun” kelimesinin sessiz harflerinden türetilmiştir.)

Projenin geliştirme sürecine göz atacak olursak; başlangıç noktası literatürdeki şekil tanıma algoritmalarının taranması olmuştur. Optik karakter tanıma (OCR) algoritmalarının kullanılmamasına; optik bilgiden ziyade dokunuşların konum – zaman bilgilerine göre bir şekil ve hareket tanıma sağlanmasına karar verilmiştir. OCR’da herhangi bir şeklin nasıl çizildiği önemli değildir, şeklin optik bilgisinden yola çıkılarak tanıma yapılır. Hal böyleyken bir düşey çubuk yukarıdan-aşağıya ya da aşağıdan-yukarıya farklı şekillerde çizildiğinde tanıma aşamasında herhangi bir ayırım yoktur. Ortaya konulmak istenen modelde ise bu ayırımın sağlanması, ayrıca belirli görsel şekli olmayan hareketlerin de (sürekli olarak yuvarlak çizmek, silme hareketi benzeri sürekli ve hızla parmağı sağala sola kaydırmak gibi) tanınması istenmektedir.

Uygun bir algoritmanın seçiminde göz önüne alınan diğer etkenler ise hız ve tanıma başarımıdır ve bu iki etken de çoğu zaman birbiriyle çelişmektedir. İlk olarak incelenen “Eğri eşleme” (Curve Matching) tabanlı algoritmalarının yüksek başarıyla tanıma gerçekleştirebildikleri gözlenmiştir. Eğri eşlemede; eğriler üzerinde eşit aralıklarla örneklenen noktaların eğrilik değerleri kıyaslanır. Bir noktanın eğrilik değeri; komşusu olan n noktaya göre konumu ile parametrik olarak belirlenen bir değerdir. Öte yandan bu yöntemin mobil cihazlarda gerçek zamanlı kullanılacak performansla uzak olduğu sonucuna varılmıştır. Burada tek önemli etken; mobil cihazların işlemci ve hafıza yönünden kısıtlı olmaları değildir. Bu cihazların çok büyük kısmında kayan noktalı işlemci de bulunmamakta, kayan noktalı işlemler emüle edildiğinden işlemci zamanı açısından çok pahalı hale gelmektedir. Bu yüzden; kayan noktalı işlemlere gereksinimi olan algoritmalar tamsayı kullanan algoritmalara kıyasla çok daha düşük performans vermektedirler.

Şekil tanıma için “katar eşleme” tabanlı algoritmalar da kullanılmaktadır ve ikinci olarak böyle bir algoritma gerçekleştirilerek denenmiştir. Bu tip algoritmaları kısaca açıklamak gerekirse; eğrinin örneklenen noktalar arasındaki yönlü doğru parçalarından oluştuğu düşünülür ve her yön bir sayısal değere denk düşmek üzere şekiller yön katarları halinde ifade edilirler. İki eğri eşlenirken; bu yön katarları eşlenerek katarlar arasında bir maliyet bulunur. Böyle bir algoritma şekil tanıma isabetinden ödün verse de hızlıdır ve büyük oranda tamsayı işlemleri ile gerçekleştirilebilir. Basit katar eşlemenin zaman kompleksitesi $O(m*n)$ (m ve n karşılaştırılan katarlardaki simge sayısı) dir. Projenin geliştirilmesi esnasında ilk olarak basit katar eşleme algoritması ile yola çıkılmış; daha sonradan katar eşlemeyi temel alan ama daha da sadeleştirilmiş $O(m+n)$ kompleksitesinde bir algoritma ortaya konmuştur. Bu algoritma şekil tanıma için basit katar eşleme ile yakın başarıya sahip olmasına rağmen çok daha hızlı çalışmaktadır. İlerleyen bölümlerde hem yukarıda anlatılan diğer algoritmalar hem de sonuç itibari ile ortaya konan bu algoritma detaylı bir şekilde incelenecektir.

Şekil tanıma algoritmasına karar verildikten sonra; projenin geliştirildiği Android Platformu’nu tanıma amacı ile çalışmalara başlanmış, geliştirme ortamı, SDK ve emülatörler kurularak Android program modeli üzerine çalışılmıştır. Projenin ikinci ayağı; Android Platformu’nun standart kullanıcı arayüzü bileşenlerinden (metin kutusu, buton, seçme kutusu vb.) şekil tanıma desteği olan alt sınıflar türetilmesidir. Buna basit bir örnek olarak tik (doğru işareti) ve çarpı (yanlış işareti) tanıyan bir seçme kutusu (checkbox verilebilir). Bilindiği üzere standart seçme kutuları üzerine tıkladığında bir kutunun seçili ya da boş hale gelmesi şeklinde çalışmaktadır. Kullanıcının tıklamak yerine bu bileşen üzerine gerçek bir tik veya çarpı işareti çizmesi ise daha doğal bir arayüz kullanım modeli olarak ortaya konulabilir.

Takip eden aşamada; türetilen şekil tanıma özelliğine sahip arayüz bileşenleri kullanılarak örnek bir uygulama ortaya konmuştur. Uygulama; buton ve menülerden ziyade; projede

hedeflenen şekil ve hareket tanıma ile kontrol edilen bir ezber kartı uygulamasıdır. Bu uygulama kullanıcıya ezber kartları yaratma, çalışma, depolama gibi imkanlar sunmaktadır ve örnek olmasına rağmen tamamen çalışır ve kullanılabilir durumdadır. Şekil ve hareketlerin kullanılmasına bir örnek olması açısından; kullanıcı bir kartın içeriğini değiştirmek istediğinde o kartın üzerine bir “e” çizerek edit ekranını açabilmektedir. Şekillerin kullanılmadığı bir arayüz modelinde ise böyle bir işlev ya ekranda fazladan yer kaplayan bir butonla, ya bir menüye girilerek ya da bu fonksiyon bir tuşa atanarak sağlanabilecektir. Projenin ADC yarışmasında tanıtımı amacı ile hazırlanan videolar izlenirse önerilen model daha iyi anlaşılacaktır:

Örnek ezber kartı programı, şekiller ile kontrol edilen kullanıcı arayüzleri için bir video:

http://www.youtube.com/watch?v=JxwGMtNS_YA

Projenin el yazısı tanıma konusundaki yeterliliğinin değerlendirilebileceği video:

<http://www.youtube.com/watch?v=C7A6OHWFsOE>

Son olarak; kod yapısına özetle göz atacak olursak yapının 4 temel kısımdan oluştuğu söylenebilir:

- Platformdan bağımsız, şekil tanıma ile ilgili algoritma ve yöntemlerin gerçekleştirildiği tanıma modülü;
- Şekil ve hareketlerin tanımlanması, depolanması ve ekranda görüntülenmesi üzerine yardımcı araçlar
- Android platformu kullanıcı arayüzü bileşenlerinin şekil tanıma özelliği kazandırılmış hallerini içeren bir arayüz bileşeni kütüphanesi
- Kütüphanenin açıklanması ve denenmesi için kullanılmak üzere yazılmış örnek uygulamalar.

TABLE OF CONTENTS

| | |
|---|-----------|
| 1 INTRODUCTION..... | 10 |
| 2 PROJECT DEFINITIONS AND PLAN | 12 |
| 2.1 DEVELOPMENT ENVIRONMENT | 12 |
| 2.2 PROJECT PLAN..... | 12 |
| 3 THEORETIC INFORMATION..... | 14 |
| 3.1 FIRST ATTEMPT: CURVE MATCHING | 14 |
| 3.2 SECOND ATTEMPT: STRING MATCHING | 15 |
| 3.3 AN ALTERED STRING MATCHING APPROACH | 16 |
| 4 ANALYSIS AND MODELLING..... | 18 |
| 5 DESIGN, IMPLEMENTATION AND TESTING | 21 |
| 5.1 PACKAGES..... | 21 |
| 5.1.1 <i>com.soydan.dkn</i> | 22 |
| 5.1.2 <i>com.soydan.dkn.demos</i> | 22 |
| 5.1.3 <i>com.soydan.dkn.flashcards</i> | 22 |
| 5.1.4 <i>com.soydan.dkn.recognition</i> | 22 |
| 5.1.5 <i>com.soydan.dkn.tools</i> | 24 |
| 5.1.6 <i>com.soydan.dkn.utility</i> | 24 |
| 5.1.7 <i>com.soydan.dkn.widgets</i> | 24 |
| 5.2 TESTING | 24 |
| 6 EXPERIMENTAL RESULTS..... | 25 |
| 7 RESULTS AND CONCLUSION..... | 26 |
| 8 REFERENCES..... | 27 |

1 INTRODUCTION

The aim of this project is to implement an accurate and fast real-time gesture recognition library, incorporate gestures into GUI elements and suggest a gesture based user interaction model. The project mainly targets mobile devices with touch screen hardware; however it is not designed exclusively for such devices. Recognition module of the project is kept platform independent and can be used with any input source that can supply 2 dimensional coordinate data over time. In fact; the whole project has been designed and tested on a device emulator that runs on PC, thus uses mouse input rather than an actual touch screen.

There are several reasons for using the word “gesture” rather than “shape” in the project name. First of all; shape recognition is associated with optical character recognition in most cases. However; this project uses online (real-time, explained below) touch motion data for recognition, rather than optical data. Although these two are alike, they are not exactly the same. Secondly; not every gesture has a specific visual shape. An example would be moving a finger left to right in fast repetitive motions, like scratching. Although this motion does not have a definite shape it is a definite gesture, and this project aims to detect such gestures as well.

Shape / handwriting recognition has been an active area of research for quite a long time, and there are numerous academic publications on it. [1] was conducted as reference throughout the creation of this project and was a valuable source for general information and the current status of these fields. Shape recognition is considered a sub category of pattern recognition that concentrates on visual data and in most cases it is closely related to image processing. Problems domains for shape and handwriting recognition can be separated into 2 main groups, on-line and off-line recognition [2]. In offline recognition, an image is scanned in order to detect visual characters that match a pattern. Offline detection algorithms are not concerned with how and when those images have been formed, hence the name offline. Examples are most OCR (Optical character recognition) software used in converting handwritten documents into a digital format. The second problem domain includes online algorithms and recognition takes place in real-time, as the characters are being input. This type of algorithms can incorporate temporal data in addition to offline recognition; however they also have to perform fast (Real-time performance is required). Further information on this topic can be found in [2].

There are also lots of professional / commercial software on the market that address gesture and especially handwriting recognition. The first successful implementation of handwriting recognition on mobile devices is considered to be the Graffiti on Palm PDA devices. Graffiti had many successors, although most of them were native to devices that they ran on. Today however; there are also many recognition based external applications on platforms that allow them, such as Windows Mobile.

This project focuses on the use of gestures on graphical user interfaces rather than fully-fledged handwriting recognition. In most cases; handwriting recognition requires precise input which can only be supplied by a special input pen. As a result; devices that do not have special pens and are navigated by fingers are not suitable for handwriting recognition. On the other hand; it can be claimed that such devices can still benefit from gesture recognition

immensely. A good use for gestures can be using them as shortcuts, i.e. firing an application or functionality when a certain gesture has been detected.

In the light of all these; the aims of the project can be concluded as follows;

- An accurate library for gesture and shape detection will be put forth. For being able to achieve this, academic publications on gesture, character and shape recognition will be viewed, a suitable algorithm will be selected and implemented.
- Considering that this library is mainly designed for mobile devices; it will be optimized to work with low processor and memory resources. A balance between recognition accuracy and execution speed will be sought.
- A gesture based user interface navigation model will be implemented. Sample applications that use this new model will be written and tested.
- Natural gestures for some common user interface functionality will be suggested. (Such as increasing the volume while a continuous clockwise circle is being drawn)

As a final note for the introduction section; I can claim that I have been successful on achieving my initial aims to a reasonable degree. I have been able to implement the library and also a sample application on Android platform which can be used for testing the usability of a gesture based user interfaces. In addition; it has received lots of positive feedback as well as ranking among the first quarter of participants in the Android Developers Contest.