

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**NEGATİF OLMAYAN MATRİS AYRIŞTIRMA YÖNTEMİ İLE
VIDEO PARMAK İZİ ÇIKARIMI**

**YÜKSEK LİSANS TEZİ
Özgün ÇIRAKMAN**

Anabilim Dalı : Elektronik ve Haberleşme Mühendisliği

Programı : Telekomünikasyon Mühendisliği

Tez Danışmanı: Prof. Dr. Bilge GÜNSEL

AĞUSTOS 2010

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**NEGATİF OLMAYAN MATRİS AYRIŞTIRMA YÖNTEMİ İLE
VIDEO PARMAK İZİ ÇIKARIMI**

**YÜKSEK LİSANS TEZİ
Özgün ÇIRAKMAN
(504071332)**

Tezin Enstitüye Verildiği Tarih : 7 Mayıs 2010

Tezin Savunulduğu Tarih : 6 Ağustos 2010

**Tez Danışmanı : Prof. Dr. Bilge GÜNSEL
Diğer Jüri Üyeleri : Doç. Dr. Neslihan Serap ŞENGÖR
Yrd. Doç. Dr. Mustafa Ersel KAMAŞAK**

AĞUSTOS 2010

Destegini benden hi esirgemeyen aileme,

ÖNSÖZ

Yüksek lisans hayatım süresince her türlü bilgisini benimle paylaşan danışmanım Prof. Dr. Bilge GÜNSEL ve gerek tez gerekse projelerde yardımını esirgemeyen sayın Doç. Dr. Neslihan Serap ŞENGÖR'e ve Yrd. Doç. Dr. Mustafa Ersel KAMAŞAK'a çok teşekkür ederim.

Hayatım boyunca desteğini benden esirgemeyen aileme ve başta projelerde birlikte çalıştığım Yiğitcan SAVRAN, Ozan GÜRSOY ve Sezer KUTLUK olmak üzere yüksek lisans hayatım boyunca bulunduğum Çoğul Ortam Sinyal İşleme ve Örüntü Tanıma Laboratuvarı'ndaki bütün arkadaşlarıma gerek her konuda yardımları gerekse okulda ve okul dışında geçirdiğim günleri güzelleştirdikleri için çok teşekkür ederim.

Bu tez çalışması TÜBİTAK tarafından 109E063 nolu "Negatif Olmayan Matris Faktörizasyonu ile Video Parmak izi Çıkarma ve Video İçerik Modelleme" isimli proje kapsamında desteklenmiştir.

Ağustos 2010

Özgün ÇIRAKMAN
(Bilgisayar Mühendisi)

İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	v
KISALTMALAR	ix
ÇİZELGE LİSTESİ.....	xi
ŞEKİL LİSTESİ.....	xiii
ÖZET.....	xv
SUMMARY	xvii
1. GİRİŞ	1
2. NEGATİF OLMAYAN MATRİS AYRIŞTIRMA	5
3. NOMA TABANLI VIDEO PARMAK İZİ ÇIKARIMI.....	11
3.1 Anahtar-Çerçeve Tabanlı Video Parmak İzi Çıkarımı	11
3.1.1 Parmak İzi Çıkarımı	12
3.1.2 İkili Hipotez Testi ve Eğitim Aşaması	14
3.1.3 Dönüşümle Değişmeyen Negatif Olmayan Matris Ayrıştırma	16
3.2 İçerik Tabanlı Video Parmak İzi Çıkarımı	21
4. TESTLER	27
4.1 Başarım Kriterleri.....	27
4.2 Anahtar-Çerçeve Tabanlı Parmak izi Sistemi için Test Sonuçları	30
4.3 İçerik Tabanlı Video Parmak izi Sistemi için Test Sonuçları	33
5. SONUÇ VE ÖNERİLER.....	39
KAYNAKLAR	43

KISALTMALAR

DCR	: Kopya Tespit Başarım Oranı
NOMA	: Negatif Olmayan Matris Ayırıştırma (Non-Negative Matrix Factorization, NMF)
DD-NOMA	: Dönüşümle Değişmeyen Negatif Olmayan Matris Ayırıştırma (Transform-Invariant Non-negative Matrix Factorization, T-NMF)
P2P	: Noktadan Noktaya (Peer to Peer)
TBA	: Temel Bileşenler Analizi (Principal Component Analysis, PCA)
2B-YTBA	: İki Boyutlu Yönlendirilmiş Temel Bileşenler Analizi (2D-Oriented Principal Component Analysis, 2D-OPCA)
TDA	: Tekil Değer Ayırıştırma (Singular Value Decomposition, SVD)
TRECVID	: TREC Video Retrieval Evaluation
BER	: Bit Hata Oranı (Bit Error Rate)

ÇİZELGE LİSTESİ

	<u>Sayfa</u>
Çizelge 4.1 : Anahtar-çerçeve tabanlı parmak izi yöntemine ilişkin eşleme uzaklıkları.....	32
Çizelge 4.2 : Anahtar-çerçeve tabanlı parmak izi sisteminin global saldırılara karşı başarımı.....	33
Çizelge 4.3 : İçerik tabanlı video parmak izi sisteminin başarımı.....	36

ŞEKİL LİSTESİ

Sayfa

Şekil 1.1 : Bir video parmak izi sisteminin genel yapısı	2
Şekil 3.1 : (a) Orjinal bir video çerçevesi ve örnek geometrik ataklı çerçeveler (b) yatayda aynalama (c) düşeyde aynalama (d) aşağı öteleme (e) sola öteleme (f) sağa öteleme (g) yukarı öteleme (h) resim içinde resim (iki boyutlu yeniden boyutlandırma) (i) tek boyutta yeniden boyutlandırma (j) döndürme	17
Şekil 3.2 : Ayırıştırma kullanılan bazı dönüşüm matrislerinin (T^m) yapıları (a) yatayda aynalama (b) düşeyde aynalama (c) sola k sütun öteleme (d) sağa k sütun öteleme. E, Z ve I matrisleri sırasıyla değişim matrisi, tam sıfır matrisi ve birim matris olup, $\mathbf{I}^{\rightarrow k}$ sütunları ok yönünde k adet ötelenmiş birim matristir.	19
Şekil 3.3 : Video çerçevelerinden öznitelik matrislerinin çıkarılması	22
Şekil 3.4 : Satır vektörlerinin ağırlıklandırılması	24
Şekil 3.5 : Ortanca değere göre eşikleme ile ikili parmak izinin elde edilmesi	25
Şekil 4.1 : Yazılımın kullanıcı arayüzü. Sol üstte bozulmaya uğramış sorgu videosunun bir çerçevesi görünürken sol altta veritabanından eşlenen referans videodaki ilgili çerçeve görünmektedir.	28
Şekil 4.2 : (a) Orjinal video çerçevesi ve TRECVID testlerinde kullanılan ataklar (b) kontrast değişimi (c) bulanıklaştırma (d) yeniden kodlama (e) gürültü ekleme (f) gama bozulması (g) resim-içi-resim (h) örüntü ekleme (i) aynalama (j) kaydırma (k) kırpma (l) yeniden boyutlandırma	29
Şekil 4.3 : Global saldırılar için BER.....	35
Şekil 4.4 : Geometrik saldırılar için BER.....	35
Şekil 4.5 : Bütün saldırılar için BER.....	36

NEGATİF OLMAYAN MATRİS AYRIŞTIRMA YÖNTEMİYLE VİDEO PARMAC İZİ ÇIKARIMI

ÖZET

Yayın kanallarında içerik izleme uygulamalarından P2P ağlarda filtrelemeye veya kopya belirleme amacıyla içeriğin takip edilmesine kadar pek çok alanda sürekli artan miktarlarda içeriğin kullanımda olması ve içerik özniteliklerinin tanımlanmasında var olan kısıtlar nedeniyle video verisini modellemek ve video içeriğine otomatik erişmek bir takım teknik sorunları da beraberinde getirmektedir. Video parmak izi sistemleri işte bu sıkıntıları ortadan kaldırmak için önerilmiş ve günümüzde pek çok uygulama alanında kullanılmaktadırlar.

Bu tez kapsamında video parmak izi sistemlerinin özellikle videoda kopya sezme alanında kullanımı üzerinde durulmuştur. Kopya sezme problemlerinde bir parmak izi sisteminin, aynı video içeriğinin kopyasını bozulmaya uğramış olsa dahi sezebilmesi istenir. Bu yüzden video parmak izi sistemleri video verisine uygulanabilecek global bozulmalara (sıkıştırma, çerçeve hızını değiştirme, gürültü, gama bozulması vb.) ve geometrik dönüşümlere (aynalama, kaydırma, kırpm, yeniden boyutlandırma gibi) karşı dayanıklı olmalıdır. Pratikte kopya sezme alanında kullanılan arama-getirme sistemleri, kendilerine verilen bir videonun, daha önceden parmak izi çıkarma uygulaması tarafından veritabanına kaydedilmiş videolar arasında olup olmadığını anlamaya çalışır.

Bu çalışmada biri anahtar-çerçeve, diğeri video tabanlı olmak üzere iki video parmak izi çıkarma yöntemi önerilmiştir. Önerilen yöntemlerin her ikisinde de video çerçevelerinden negatif olmayan matris ayrıştırma (NOMA) yöntemi ile elde edilen baz ve kodlama matrisleri kullanılarak elde edilen video parmak izi vektörlerinin kullanımı önerilmektedir.

Anahtar-çerçeve tabanlı birinci yöntemde video içeriğinin tamamı değil, sahnelerden veya belirli süre aralıklarında mümkün olduğunda az değişen video parçalarından alınan çerçeveler kullanılır. Buna göre önce, alınan çerçevelerin boyutlarını düşürmek amacı ile DC katsayıları kullanılarak görüntü yeniden boyutlandırılır. Ardından yeniden boyutlanmış olan görüntü NOMA ile baz ve kodlama matrislerine ayrıştırılarak video çerçevesine ilişkin öznitelikler elde edilir. Ardından hem öznitelik boyutlarını azaltmak hem de arama-getirme esnasında kullanılacak benzerlik ölçütünü kolayca hesaplayabilmek için bu öznitelik vektörleri ayrı ayrı Gauss vektörleri ile çarpılarak ağırlıklandırılır ve anahtar-çerçeve için parmak izi elde edilir. Bu yöntemde arama-getirme esnasında eşleme için Öklid metriği kullanılmıştır. Önerilen bu ilk yöntemin global bozulmalarda iyi çalıştığı gözlenirken geometrik dönüşümlerde sistemin performansının artırılması açısından, çıkarılan parmak izi ile eşleme sağlanamaması halinde literatürde bulunan DD-NOMA [1] yöntemi ile de desteklenmesi önerilmektedir.

Video tabanlı ikinci yöntemde ilk yöntemden farklı olarak parmak izleri tek bir video çerçevesi için değil, birbirini takip eden bir grup video çerçevesi için çıkarılır. Şayet

parmak izleri video çerçevelerinden hiçbirini dışarıda bırakmayacak şekilde birbirini takip eden parçalar halinde çıkarılırsa bir video klibi içerisinde herhangi bir noktaya rastgele erişmek olanaklı hale gelir. Bu yöntemde bir video çerçevesi önce alt bloklara ayrılır. Ardından hem çerçevenin kendisi hem de alt bloklar DC katsayıları kullanılarak yeniden boyutlandırılır. Birbirini takip eden çerçevelere ait bloklar her blok tipi için ayrı ayrı veri matrislerine sütun vektörleri olarak dizilir. Böylelikle parmak izi çıkarılacak süreye karşı gelen çerçeve adedi uzunluğunda, blok tipi kadar farklı veri matrisi elde edilir. Elde edilen bu veri matrisleri NOMA yöntemi ile ayrıştırılır ve blok tipi kadar baz ve kodlama matrisi elde edilir. Elde edilen bu matrisler ilk yöntemde olduğu gibi Gauss ile ağırlıklandırılarak her blok tipi için kayan noktalı sayı tipinden bir parmak izi elde edilir. Burada elde edilen parmak izleri ortanca değere göre eşikleme yöntemi ile ikili sayı tipine dönüştürülebilir. Elde edilen parmak izlerinin birleşiminden de video parçasını temsil eden son parmak izi oluşturulur.

Performans testleri TRECVID [2] 2002, 2009 ve 2010 veritabanlarından alınan video klipleri üzerinden gerçekleştirilmiştir ve başarımlar doğruluk (precision), hatırlama (recall) ve bunların harmonik ortalaması olan kopya tespit başarımlar oranı (DCR) cinsinden raporlanmıştır.

VIDEO FINGERPRINTING VIA NON-NEGATIVE MATRIX FACTORIZATIONS

SUMMARY

Modelling and retrieving video content is a difficult problem due to increasing amounts of media in applications such as broadcast monitoring, P2P network filtering or copy detection and constraints on media content representations. In order to deal with these problems, video fingerprinting systems have been proposed and used in wide range of applications.

Copy detection features of video fingerprinting systems are the main concern of this work. In copy detection applications, a fingerprinting system should be able to match an input query video clip with its original content which has been previously introduced to the system. Thus, video fingerprinting system should be robust against possible modifications on input data. These modifications are generally mentioned under two different classes as global distortions and geometric transformations. Video re-encoding, frame rate conversion, changes in brightness, contrast and gamma, introduction of additive noise and blur can be given as examples of global distortions while flipping, shifting, cropping and resizing attacks are the examples of geometric transformations.

In this work two video fingerprinting approaches are proposed. Both approaches use non-negative matrix factorization (NMF) as feature extraction tool but differs especially in selection of the input data. In the first approach, extracted fingerprints only correspond to the selected key-frames of the input video clip, while in the second method fingerprints are extracted for groups of successive frames.

In the first approach, each fingerprint is generated for individual key-frames which can be selected from scenes or certain periods of the video clip. Selected key-frames are resized using DC coefficients in order to shrink data size. Then each resized frame is factorized into two feature matrices using NMF. These feature matrices are used while forming the final fingerprint. Elements of the final fingerprint are calculated as the feature matrices which are further weighted by using different pseudo-random Gaussian vectors for each element. This weighting process helps to decrease the fingerprint size and provides a convenient way to find fingerprint matching criterion in order to be used in the search & retrieval phase. Similarity metric selected for the fingerprints is the Euclidean metric because of its simplicity and efficiency. While the proposed system is shown to be robust against global distortions, it needs to be supported by an additional step in order to increase robustness against geometric transformations. Thus, when a proper match cannot be found, transform invariant non-negative matrix factorization (T-NMF) [1] is used before terminating the search & retrieval process to make a decision.

In the second approach, video fingerprints are extracted from a sequentially gathered group of frames, rather than individual key-frames as in the first method. By selecting frame groups (without leaving any single frame out), video fingerprints can

now represent any single point in the video clip and random access on video data will be possible. The first step in this approach –is to obtain sub-image-blocks from each input frame into. Then, multiple (=total number of block types and frame itself) data matrices are created. Each data matrix is composed by corresponding sequences of sub-blocks or frames themselves. These prepared data matrices are then factorized using NMF and feature matrices for each block type are obtained. Feature matrices are then weighted with pseudo-random Gaussian vectors as in the first approach and floating point type sub-fingerprints are obtained for each block type. These floating point sub-fingerprints can then be transformed into binary sub-fingerprints by applying median thresholding on each of them individually. Final fingerprint is obtained by merging sub-fingerprints which represent different block types.

Proposed systems are tested on TRECVID [2] 2002, 2009 and 2010 video databases and performance of the system is evaluated by using precision and recall rates as well as their harmonical average DCR (detection cost rate).

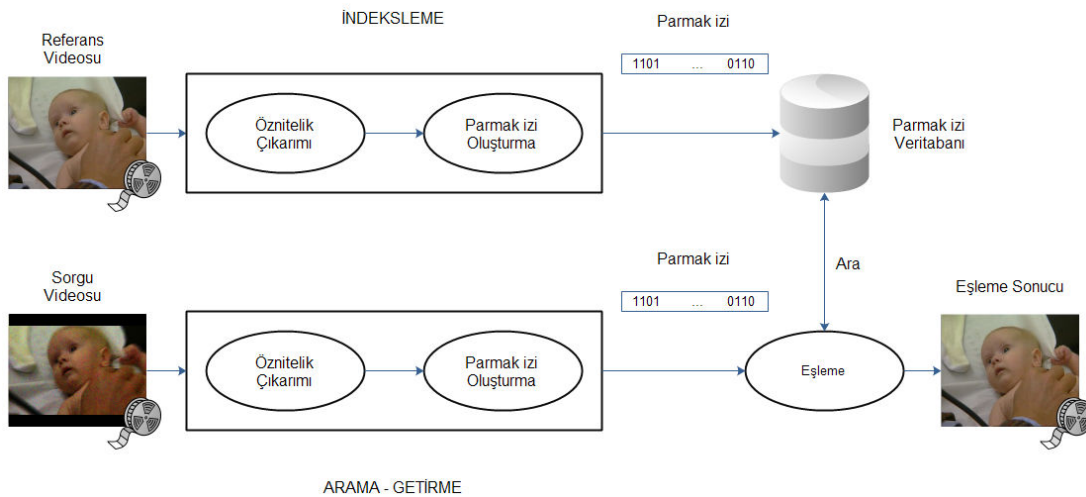
1. GİRİŞ

Yayın kanallarında içerik izleme uygulamalarından P2P ağlarda filtrelemeye, kopya içeriğin takip edilmesinden benzer içeriklerin tespitine kadar pek çok alanda sürekli artan miktarlarda video içeriğinin kullanımda olması ve içerik özniteliklerinin tanımlanmasında var olan kısıtlar nedeniyle video verisini modellemek ve video verisine otomatik erişmek bir takım teknik sorunları da beraberinde getirmektedir. Video parmak izi sistemleri işte bu sıkıntıları ortadan kaldırmak için önerilmiş ve günümüzde pek çok uygulama alanında kullanılmaktadırlar. Bir video parmakizi sisteminin, video kliplerinin içeriklerini yegane olarak belirleyebilmesi, videoların içerisinden onları algısal olarak temsil edebilecek ayırd edici öznitelikleri çıkarabilmesi gereklidir.

Bu çalışmada amaç kopya algılamada kullanılabilecek bir video parmak izi çıkarma yöntemi geliştirmektir. Kopya algılamada kullanılacak olan bir parmakizi çıkarma sisteminin videonun algısal içeriğini değiştirmeyen dönüşümlere dayanıklıyken yeganeliği de koruması gerekir. Algısal dönüşümler örüntü ekleme, bulanıklaştırma, gama değişimi, gürültü ekleme, yeniden boyutlandırma gibi video verisini sayısal olarak değiştirse de kendisini izleyen kişinin videoyu algılamasında farklılık yaratmayacak dönüşümler olarak tanımlanabilir. Hedef, bu şekilde bir bozulmaya uğramış olan bir video klip ya da klip içerisinden alınan daha ufak bir bölümden çıkarılan parmakizinin, orjinal video klibinden çıkarılan parmakizi ile olabildiğince yakın olması ve bir arama getirme işlemi sonucunda bu ikililerin doğru olarak eşlenebilmeleridir.

Bu noktada video parmak izi çıkarma ve arama-getirme sistemlerinden bahsetmek faydalı olacaktır. Bir video parmak izi sisteminde yapılan işlemleri, indeksleme ve arama-getirme sistemleri olarak ikiye ayırmak mümkündür. İndeksleme sistemi, Şekil 1.1’de üst dalda görülebileceği üzere, sisteme sunulan videoların çeşitli yöntemlerle parmak izlerini çıkarıp, elde edilen parmak izlerini bir veri tabanına kaydetmelidir. Arama-getirme sistemi ise sisteme sunulan bir video klibinin veritabanında önceden var olan videolardan biri veya birkaçıyla uyuşup

uyuşmadığını, uyuşuyorsa bu uyuşmanın söz konusu videolar içindeki konumunu ve miktarını tespit etmelidir. (bkz. Şekil 1.1 alt dal). Arama-getirme sistemi bu işlevi gerçeklemek için sisteme sunulan sorgu videosuna ait parmak izlerini (indekslemede kullanılan aynı yöntemle) elde eder ve elde edilen bu parmak izlerini önceden belirlenen bir eşleştirme kriterine göre veritabanında arar. Bir örnekle anlatmak gerekirse, 1 saniyelik çözünürlüğe sahip bir parmak izi sistemi, verilen videoları birer saniyelik parçalara ayırıp her parçaya karşı düşen bir parmak izi çıkarır. Çıkarılan bu parmak izleri veritabanına kaydedilir. Böyle bir sistemde sorguda kullanılan videolar da aynı şekilde birer saniyelik alt parçalara ayrılıp parmak izleri elde edilir ve bu parmak izleri veritabanındaki parmak izleri arasında aranır. Burada dikkat edilmesi gereken nokta, parmak izleri birer saniyelik parçalar halinde çıkarıldığı için sistem bir saniyeden daha küçük bir eşleme ya da bir saniyeden daha hassas konumlandırma yapamaz. Buna da sistemin çözünürlüğü (*granularity*) denir.



Şekil 1.1 : Bir video parmak izi sisteminin genel yapısı

Söz konusu sistemlerde kullanılan parmak izi çıkarma tekniklerine detaylı olarak göz atılırsa, mevcut parmak izi çıkarım yöntemlerinin çözünürlük, parmak izi veri tipi, çalışılan dönüşüm uzayı, ilgilenilen bozulma türleri veya video özniteliklerinin çıkarılma yöntemi gibi konularda farklılık gösterdiği görülebilir. Bu yöntemlerde parmak izlerinin çıkarımı genel olarak ya anahtar video çerçeveleri ya da videonun alt parçaları için bir çerçeve grubu kullanılarak gerçekleştirilmektedir.

Burada kısaca literatürde mevcut parmak izi çıkarma tekniklerine değinilecektir . Dönüşüm uzayında çalışan yöntemlerden biri olarak, Lee ve Yoo (2008) tarafından 2B yönlendirilmiş temel bileşenler analizi “2B-YTBA” (*2D - Oriented Principal*

Component Analysis, 2D-OPCA) kullanarak video anahtar çerçevelerinin ılgın eşdeğişkin alanlarını (*affine covariant regions*) bulmaya dayalı bir yöntem önerilmiştir [3]. Bu yöntem, eğitim sırasında hesaba katılan global saldırılara karşı gürbüz olsa da, geometrik saldırılara karşı gürbüzlüğü çıkarılan ılgın alanların etkinliğine bağılı olduğundan parmak izinin önemli geometrik bozulmalara karşı dayanıklılığı garanti edilememektedir.

Bir başka yöntem de Massoudi ve Lefebvre (2006) tarafından önerilen, video anahtar çerçeveleri içerisinde belirlenen öznitelik noktaları civarından açısız kısımlama (RASH) kullanılarak elde edilen gradyan yönelimlerinin histogramları kullanılarak parmakizinin elde edilmesidir [4]. Bu yöntem, yüksek hesaplama yüküne karşık global saldırılara karşı oldukça yüksek bir gürbüzlüğe sahiptir. Ancak, bu yöntem de [3]'de olduğu gibi geometrik saldırılara karşı gürbüzlüğü garanti edememektedir.

Önerilen çoğu parmak izi çıkarım yöntemi temel olarak global bozulmalara karşı etkin olsa da literatürde geometrik dönüşümlere karşı dayanıklı parmak izleri için de çalışmalar vardır. Video çerçevelerini eşmerkezli dairesel parçalara ayırıp her bir parçadan elde edilen matrisler üzerinde tekil değer ayrıştırması “TDA” (*Singular Value Decomposition, SVD*) yapılarak elde edilen özniteliklerden parmak izleri çıkarmak geometrik dönüşümlere karşı gürbüz olduğu iddia edilen yöntemlerden biridir [5]. Bunun haricinde video çerçeveleri içinden alınan eşmerkezli dairesel parçalardan çıkarılan moment değişmezleri ile ardışıl çerçeveler arasındaki farkları kullanmak da yine aynı kişiler tarafından önerilmiş, hafif geometrik dönüşümlere ve özellikle döndürmeye karşı da başarılı oldukları gösterilmiştir [5].

Literatürde geometrik dönüşümlere çözüm bulmayı hedefleyen diğeri bir yöntem de [1]'de önerilen dönüşümle değişmeyen negatif olmayan matris ayrıştırma “DD-NOMA” (*Transform Invariant Non-negative Matrix Factorization, T-NMF*) yöntemidir. Burada dayanıklı olması hedeflenen saldırı türleri önceden belirlenir ve ilgili dönüşüm matrisleri NOMA'da hesaba katılır. Bu yöntem şayet olası saldırılar önceden biliniyorsa oldukça etkili olmakla birlikte pratikte uygulanabilecek bütün saldırılar bilinemeyeceğinden ve yöntem tam anlamıyla bir parmak izi üretmediğinden yine de problemi tam olarak çözememektedir.

Bu çalışma kapsamında biri tek tek anahtar-çerçeveler diğeri video çerçeve grupları üzerinde işlem yapan, negatif olmayan matris ayrıştırma yöntemi “NOMA” (*Non-*

Negative Matrix Factorization, NMF) tabanlı iki algısal parmakizi çıkarma yöntemi gerçekleştirilmiştir. NOMA yönteminin etkisi veriyi yerel özniteliklerini vurgulamayı sağlayan seyrek, parça-tabanlı temsil edebilmesindendir [6]. NOMA özellikle son zamanlarda görüntü kıyımlama [7], video temsil etme [8] ve video indeksleme [1] gibi görüntü ve video analizi uygulamalarında giderek artan bir sıklıkla uygulanmaktadır.

Bu tez çalışmasında önerilen yöntemler parmak izlerinin yeganeliği ve algısal saldırılara karşı gürbüzlüğü yönünden test edilmiştir. Sonuçlar TRECVID [2] veritabanından alınan örnek videolar üzerinden verilmiş olup başarımlar kriterlerinde de TRECVID baz alınmıştır.

2. NEGATİF OLMAYAN MATRİS AYRIŞTIRMA

Video, görüntü veya ses içerikleri gibi büyük boyutlu veriler üzerinde yapılan çalışmalarda, veri içeriğindeki görece önemsiz bileşenleri arkaplana atıp ele alınan problemin çözümünde daha işe yarar bileşenlerin ortaya çıkarılması önem kazanır. Pratikte boyut düşürme, gerek maliyetleri düşürmek gerek içerik temsilini yalınlaştırarak başarımı artırmak açısından önemlidir [9]. Bunun için literatürde bulunan bir çok ayrıştırma yöntemi veri boyutunu küçültüp önemli bileşenleri ortaya çıkarmaya çalışır. Bu tür ayrıştırma yöntemlerinin en bilinenleri olarak tekil değer ayrıştırması “TDA” (*singular value decomposition, SVD*) ve onun ortak değişinti matrisi üzerinde uygulanmasından doğan temel bileşenler analizi “TBA” (*principal component analysis, PCA*) gösterilebilir [10]. Negatif olmayan matris ayrıştırma “NOMA” (*Non-negative Matrix Factorization, NMF*) yöntemi de ardışık bir matris ayrıştırma yöntemi olarak literatüre eklenmiş ve özellikle son yıllarda, yüz tanıma [11], nota tanıma [12], kaynak ayrıştırma [13], görüntü kısımlama [7,14] ve video parmak izi çıkarımı gibi alanlarda artan bir kullanım alanı bulmaya başlamıştır [11-23]. NOMA doğası gereği zaten negatif elemanlar içermeyen bu tür veriler için oldukça uygun bir yöntem olarak değerlendirilir [9].

Çok daha önce önerilmiş ve kullanılmış olsa da [24,25] NOMA yönteminin yaygınlaşması Lee ve Seung’un 1999 yılı çalışması ile birlikte olmuştur [9]. Temelde NOMA, bir veri matrisini biri karışım (veya baz) diğeri kodlama matrisleri olmak üzere iki adet negatif olmayan elemanlardan oluşan çarpana ayırır ve veri boyutunun düşürülmesini sağlar. Alternatiflerinin aksine, barındırdığı negatif olmama koşulu sayesinde işlem sonunda elde edilen çarpanların kazandığı toplanabilirlik özelliği ile verilerin parça tabanlı gösterimini mümkün kılması önemli özelliğidir. Lee ve Seung NOMA’nın iteratif adımlarında kullanılmak üzere baz ve kodlama matrisleri için çarpımsal güncelleme kuralları önermiş ve böylelikle yöntemin kullanım kolaylığını artırırken yaygınlığını da artırmıştır [9]. Buna ek olarak NOMA maliyet fonksiyonuna kodlama matrisini üzerinden konulan ek terimler sayesinde ayrıştırma sonucunda elde edilen çarpanların seyrekliğinin (*sparseness*) artması sağlanmış,

böylelikle elde edilen baz matrisinin içerik verisini parça tabanlı temsil etme özelliği pekiştirilmiştir [13][16-19].

Buradan itibaren NOMA yöntemi ile ilgili matematiksel ifadelere bu çalışmada kullanıldığı kapsamda değinilecektir. Buna göre, kullanılacak veri $n \times m$ boyutlu bir V matrisi ile temsil edilecektir. Denklem (2.1)'de görüldüğü üzere $V \in R^{n \times m}$ matrisi, NOMA yöntemi kullanılarak belirli bir yaklaşıklıkla iki çarpana ayrıştırılır [15].

$$V \approx WH \quad (2.1)$$

Burada $W \in R^{n \times r}$ baz (karışım) matrisi iken $H \in R^{r \times m}$ kodlama matrisi olup r NOMA yönteminin ayrıştırma düzeyini (*rank*) belirleyen bir parametredir. Rank, ayrıştırma matrislerinin boyutunu belirlediğinden NOMA'nın boyut küçültme seviyesini belirler. Rank küçüldükçe boyut daha fazla düşürülebilirken ayrıştırma işleminin yaklaşıklık hatası ise büyür. Bu nedenle ayrışım düzeyi ilgilenilen probleme göre dikkatle seçilmelidir.

NOMA yönteminin gerçekleşmesinde literatürde birden fazla maliyet fonksiyonu kullanılıyorsa da [7,17,18] basitliği ve bu basitliğin karşılığında önemli bir kayba sebebiyet vermemesi ile en çok tercih edilen maliyet fonksiyonu denklem (2.2) ile verilen, ortalama karesel geri çatma hatasıdır.

$$F = \frac{1}{2} \|V - WH\|^2 = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^m \left((V)_{ij} - (WH)_{ij} \right)^2 \quad (2.2)$$

Burada i ve j alt indisleri sırasıyla ilgili matrislerin satır ve sütunlarını belirtir. Karesel hata fonksiyonunun, W ve H matrislerine ayrı ayrı konveks olması nedeniyle maliyet fonksiyonu gradyan inişi yöntemi ile eniyilendiğinde W ve H matrisleri için ayrı ayrı çarpımsal güncelleme kuralları bulunabilir [15].

Gradyan inişi yöntemine göre H matrisi için güncelleme bağıntısı (2.3)'deki gibi ifade edilebilir. Burada $j=1, \dots, m$ olup H matrisinin sütunlarını, $a= 1, \dots, r$ olup H matrisinin satırlarını belirtir. η_{aj} gradyan inişindeki kullanılan adım boyunu gösterir. [9].

$$(H)_{aj} \leftarrow (H)_{aj} - \eta_{aj} \frac{\partial F}{\partial (H)_{aj}} \quad (2.3)$$

Burada görüldüğü gibi güncelleme kuralı henüz çarpımsal bir formda değildir. Bu güncelleme adımını çarpımsal hale dönüştürebilmek için (2.3)'teki kısmi türev (2.4) eşitliğindeki gibi açılabilir.

$$\begin{aligned} \frac{\partial F}{\partial (H)_{aj}} &= \frac{\partial}{\partial (H)_{aj}} \left(\frac{1}{2} \sum_{i=1}^n \sum_{k=1}^m ((V)_{ik} - (WH)_{ik})^2 \right) \\ &= \left(\frac{1}{2} \sum_{i=1}^n 2((V)_{ij} - (WH)_{ij}) \frac{\partial}{\partial (H)_{aj}} \sum_{k=1}^m ((V)_{ik} - (WH)_{ik}) \right) \\ &= - \sum_{i=1}^n ((V)_{ij} - (WH)_{ij}) (W)_{ia} \\ &= - (W^T V)_{aj} + (W^T WH)_{aj} \end{aligned} \quad (2.4)$$

Burada bulunan kısmi türev eşitliğine göre, adım boyu (2.5)'deki gibi seçildiği takdirde H için çarpımsal güncelleme kuralı (2.6)'da verildiği şekilde bulunabilir [9].

$$\eta_{aj} = \frac{(H)_{aj}}{(W^T WH)_{aj}} \quad (2.5)$$

$$(H)_{aj} \leftarrow (H)_{aj} \frac{(W^T V)_{aj}}{(W^T WH)_{aj}} \quad (2.6)$$

Aynı şekilde, W matrisine ilişkin için güncelleme kuralı (2.7)'deki gibi ifade edilebilir. Burada $i=1, \dots, n$ olup W matrisinin satırlarını, $a=1, \dots, r$ ise W matrisinin sütunlarını belirtir. λ_{ia} daha öncekine benzer şekilde gradyan inişi yönteminin adım boyunu gösterir [9].

$$(W)_{ia} \leftarrow (W)_{ia} - \lambda_{ia} \frac{\partial F}{\partial (W)_{ia}} \quad (2.7)$$

Çarpımsal güncelleme kuralı elde etmek için (2.7)'deki kısmi türev (2.8)'deki gibi düzenlenebilir [9].

$$\begin{aligned}
\frac{\partial F}{\partial (W)_{ia}} &= \frac{\partial}{\partial (W)_{ia}} \left(\frac{1}{2} \sum_{k=1}^n \sum_{j=1}^m \left((V)_{kj} - (WH)_{kj} \right)^2 \right) \\
&= \left(\frac{1}{2} \sum_{j=1}^m 2 \left((V)_{ij} - (WH)_{ij} \right) \frac{\partial}{\partial (W)_{ia}} \sum_{k=1}^n \left((V)_{kj} - (WH)_{kj} \right) \right) \\
&= - \sum_{j=1}^m \left((V)_{ij} - (WH)_{ij} \right) (H)_{aj} \\
&= - (VH^T)_{ia} + (WHH^T)_{ia}
\end{aligned} \tag{2.8}$$

Buradan görüldüğü üzere, adım boyu (2.9)'daki gibi seçilirse, W matrisi için de çarpımsal güncelleme kuralı (2.10)'daki gibi elde edilmiş olur.

$$\lambda_{ia} = \frac{(W)_{ia}}{(WHH^T)_{ia}} \tag{2.9}$$

$$(W)_{ia} \leftarrow (W)_{ia} \frac{(VH^T)_{ia}}{(WHH^T)_{ia}} \tag{2.10}$$

Yukarıda (2.6) ve (2.10)'da belirtilen güncelleme adımları ile ilgili önemli bir nokta bu adımların dönüşümlü olarak gerçekleşmesidir. Buna göre her bir NOMA adımında önce kodlama matrisi olan H güncellenir, ardından kodlama matrisinin aldığı yeni değerle baz (karışım) matrisi olan W güncellenir.

Ayrıca bu noktada alışılmışın dışında bir kullanım olarak gradyan inişi yönteminde kullanılan adım büyüklüklerinin değişken seçilmesi görülebilir. Adım boylarının bu şekilde seçilmesinin hem yukarıda belirtildiği gibi W ve H için güncelleme adımlarının çarpımsal forma getirilmesini, hem de maliyet fonksiyonunun değerinin NOMA adımları esnasında artmayacağını garantilenmesini sağladığı Lee ve Seung tarafından gösterilmiştir [15,26].

NOMA ile ilgili diğer bir nokta ise diğer iteratif yöntemlerde olduğu gibi W ve H için başlangıç değerlerinin belirlenmesidir [9]. Genelde NOMA yönteminde başlangıç değerlerinin belirlenmesi çok büyük bir önem taşımaya da, en iyileme esnasında mutlak en düşük noktaya erişmenin pratikte pek mümkün olmaması nedeniyle, farklı

başlangıç değerleri altında farklı ayrıştırma sonuçları alınabileceği unutulmamalıdır [9].

3. NOMA TABANLI VIDEO PARMAC İZİ ÇIKARIMI

Bölüm 1’den hatırlanacak olursa bir parmak izi sisteminin iki temel işlevi bulunur. Bu işlevler; referans video kaynağının her bir alt parçasından çıkarılan video parmak izlerinin bit dizileri halinde ya da öznitelik vektörleri olarak parmak izi tablolarına kaydedilmesi ve sorgu videosundan gelen video çerçevelerinden içerik analizi sonucu elde edilen öznitelik vektörlerinden çıkarılan parmak izi bit dizilerinin arama tablosu yardımıyla referans video ile eşlenmesidir. Bu işlevlerden ilkinde indeksleme, ikincisine de arama getirme denir (bkz. Şekil 1.1).

Bu tez çalışması kapsamında öncelikle video sahnelerine ait anahtar çerçeveler (*key frames*) için bir parmak izi çıkarım yöntemi geliştirilmiş ardından bu yöntemin kopya sezme uygulamalarında karşılaşılan geometrik bozulmalara dayanımını artırmak için literatürde önceden önerilmiş DD-NOMA yöntemi ile nasıl desteklenebileceğinden bahsedilmiştir. Son olarak da anahtar çerçeveler için önerilen parmak izi çıkarım yöntemi baz alınarak video çerçeve dizileri için kullanılabilecek bir video parmak izi çıkarma metodu tasarlanmıştır. Önerilen her iki yöntem de sırasıyla aşağıdaki alt bölümlerde anlatılmaktadır.

3.1 Anahtar-Çerçeve Tabanlı Video Parmak İzi Çıkarımı

Bu bölümde anahtar-çerçeve tabanlı bir video arama-getirme sistemi için bir parmak izi çıkarma sistemi önerilmiştir. Böyle bir sistemde parmak izleri video klipleri üzerinde sahne geçiş sezme yöntemleri kullanılarak [27,28] seçilen anahtar-çerçeveler için çıkarılır ve eşlemeler yine anahtar çerçeveler üzerinden gerçekleştirilir. Böyle bir sistemin avantajı karmaşıklığın çok daha düşük olması ve çok daha hızlı çalışabilmesi iken, sistem parmak izlerini sadece anahtar-çerçeveler için çıkardığından bu çerçeveler dışında kalan video içeriğine rastgele erişime olanak sağladığını söylemek mümkün değildir. Bu nedenle böyle bir sistemde sorgu ve referans video klipleri arasında detaylı bir konum eşlemesi gerçekleştirmek mümkün değildir. Bahsedilen problemler göz önünde bulundurularak Bölüm 3.2’de ardışıl video çerçevelerinden oluşturulan görüntü grupları için alternatif bir video parmak

izi çıkarma yöntemi önerilmiştir. Bu bölümde ise sırasıyla anahtar-çerçeve bazlı parmak izi çıkarımı için önerilen bir sistem ve önerilen sistemin eğitimi iki alt başlık altında anlatılmaktadır. Ardından takip eden bir alt başlıkta da önerilen sistemin video kopya sezme uygulamalarında sıkça karşılaşılan geometrik saldırılara karşı gürbüzlüğü artırılması için kullanılan, literatürde var olan [1] dönüşümle değişmeyen NOMA (DD-NOMA) yöntemi kullanılarak nasıl desteklenebileceği üzerinde durulmuştur.

3.1.1 Parmak İzi Çıkarımı

Önerilen parmak izi çıkarma yönteminin temeli NOMA tarafından üretilen baz (karışım) ve kodlama matrislerinin Gauss ağırlıklandırma ile daha düşük boyutlu bir uzaya indirgenmesine dayanır. Sistemin girişi bir video klibinden çıkarılan anahtar-çerçevelerdir. Sistem, gelen sorgu videosunun anahtar-çerçeveleri için parmak izleri çıkarır ve bu çıkarılan parmak izlerini veri tabanına önceden kaydedilmiş olan referans videoların anahtar-çerçevelerine ait parmak izleri ile karşılaştırır.

Yöntemde işlemsel karmaşıklığı azaltmak için video çerçevelerinin DC görüntülere çevrilerek kullanılması önerilir. Görüntü verisinin DC katsayılarının kullanılması ayrıştırma sonucu elde edilen gösterimin yeganeliğini etkilemediği gibi veri boyutunun küçülmesine dolayısıyla işlemsel karmaşıklığın azalmasına sağlar. Bu nedenle bu çalışmada işlemsel karmaşıklığın azaltılması için video parmak izleri oluşturulurken video çerçevelerinin DC katsayılarının kullanılması önerilmektedir. Buna göre, alınan bir anahtar-çerçeve $F \in R^{w \times h}$ seçilen DC katsayılarının kullanılması ile $p \times q$ olarak yeniden boyutlandırılır ve NOMA için veri matrisi $V \in R^{p \times q}$ şeklinde elde edilir.

Sırasıyla karışım ve kodlama matrisleri olan $W \in R^{p \times r}$ ve $H \in R^{r \times q}$, giriş DC matrisi $V \in R^{p \times q}$ kullanılarak NOMA yöntemi ile elde edilir. Bir sonraki adım olan Gauss ağırlık vektörleri ile çarpımda kullanılmak üzere elde edilen çarpım matrisleri satır vektörleri y^w ve y^h olarak (3.1) denklemlerindeki gibi dizilir.

$$\begin{aligned} y^w &= [W_{l,1} \quad W_{l,2} \quad \dots \quad W_{l,r}]^T \quad l = 1, 2, \dots, p \\ y^h &= [H_{1,j} \quad H_{2,j} \quad \dots \quad H_{r,j}]^T \quad j = 1, 2, \dots, q \end{aligned} \tag{3.1}$$

Burada r NOMA'nın ayrıştırma düzeyi olmak üzere y^w ve y^h vektörleri sırasıyla $p \cdot r$ ve $r \cdot q$ boyutludur.

Kayan noktalı sayı tipinden nihai parmak izi bu y^w ve y^h vektörlerinin sözde-rastlantısal Gauss vektörleri ile ağırlıklandırılmasından oluşur. Bu ağırlıklandırma işlemi y^w ve y^h için ayrı ayrı yapılır. Bunun gerekçesi parmak izi vektörünün N elemanlı bir parçasının tamamen W , kalan M elemanlı parçasının ise tamamen H 'dan oluşmasının hedeflenmesidir. Buna göre $D = N + M$ elemanlı nihai parmak izi vektörü $x \in R^D$, Gauss ağırlıklandırma vektörleri g_l^w ve g_l^h kullanılarak (3.2)'ye göre oluşturulur.

$$x_k = \begin{cases} \sum_{l=1}^{p \cdot r} y_l^w g_l^w, & k = 1, 2, \dots, N \\ \sum_{l=1}^{q \cdot r} y_l^h g_l^h, & k = (N+1), (N+2), \dots, (N+M) \end{cases} \quad (3.2)$$

Buradan görülebileceği üzere parmak izi vektörünün ilk N elemanı y^w vektörünün N adet farklı sıfır ortalamalı birim varyanslı sözde-rastlantısal Gauss vektörü ile çarpılması ile elde edilmişken parmak izi vektörünün geriye kalan M elemanı y^h vektörünün M adet farklı yine sıfır ortalamalı birim varyanslı sözde-rastlantısal Gauss vektörü ile çarpılması ile elde edilir. Bu ağırlıklandırma işleminin önemi, birbirlerine göre yaklaşık olarak dik olan NOMA baz vektörlerinin Gauss dağılımına uydurulması ile hem baz vektörlerinin bağımsızlığının sağlanmaya çalışılması hem de elde edilen parmak izinin elemanlarının dağılımının görece basit olan Gauss formuna getirilmesidir. Bunun nedeni, önerilen parmak izi sisteminde parmak izi eşleme probleminin ikili hipotez karşılaştırma problemi olarak düşünülmesi ve zaten bağımsıza yakın olan özneliliklerin bu çarpım sayesinde Gauss formuna getirilerek ikili hipotez testinin kolaylıkla gerçekleştirilmesinin hedeflenmesidir. Önerilen bu Gauss ağırlıklandırma yöntemi Monga ve Mıhçak'ın görüntü parmak izi çalışmasından esinlenilmiştir [7]. Monga ve Mıhçak bu yöntemi yine NOMA tabanlı bir parmak izi çıkarma algoritmasının içerisine eklemiştir ancak orada bahsedilen problem görüntü eşleme tabanlı olup verinin NOMA yöntemine hazırlanma biçimi ve özneliliklerin kullanımı bu çalışmaya göre farklılık göstermektedir. Söz konusu ikili hipotez testi bölüm 3.1.2 altında detaylı olarak incelenmiştir.

Burada dikkat edilmesi gereken nokta, önerilen parmak izi çıkarma yöntemi veri boyutlarını girişe verilen çerçeve boyutları olan $(w \times h)$ ’dan parmak izi vektörünün boyutu olan $(D = N + M)$ ’ye düşürmektedir. $(D = N + M) \ll (w \times h)$ olduğu düşünüldüğünde bu işlem ciddi bir boyut düşmesine karşı gelmektedir. Burada D ’nin seçimi uygulamaya bağlı olmakla birlikte önerilen sistemde N ve M 16, dolayısıyla D de 32 olarak belirlenmiştir. Parmak izi vektörünün boyutu video parmak izi veritabanının boyutunu ve arama getirme süresini etkilediğinden önemlidir.

3.1.2 İkili Hipotez Testi ve Eğitim Aşaması

Bir önceki bölümde bahsedilen ve Monga ile Mıhçak’dan [7] esinlenildiği belirtilen ve parmak izi eşlemede kullanılması önerilen ikili hipotez testini açıklamak için video klipi içerisinden alınan i . anahtar-çerçeveden çıkarılan D boyutlu parmak izi vektörüne x diyelim. Bir başka anahtar-çerçeveden çıkarılan parmak izi ise z olsun. Video üzerine uygulanan bir saldırının sonucunda parmak izi vektörüne binen gürültü n_A ile temsil edilsin. Böylelikle iki sınıflı hipotez testi; ilk hipotez H_0 , alınan parmak izi y ’nin i . anahtar-çerçevenin bozulmuş bir durumuna karşı geldiği durum ve ikinci hipotez H_1 , alınan parmak izi y ’nin i . anahtar-çerçeveden başka bir çerçeveden (örneğin j . çerçeveden veya başka bir videodan) olduğu durum şeklinde kurulabilir. Hipotezlerin ifadesi (3.3)’de görülmektedir.

$$\begin{aligned} H_0 : y &= x + n_A \\ H_1 : y &= z \end{aligned} \tag{3.3}$$

Bu eşitlik $v = y - x$ şeklinde vektörlerin farkı biçiminden de ifade edilebilir ve böylelikle hipotezlerden biri bozulma gürültüsü, diğeri de iki orijinal çerçeveden çıkarılmış parmak izlerinin farkı biçiminde ifade edilmiş olur. (bkz. denklem (3.4))

$$\begin{aligned} H_0 : v &= n_A \\ H_1 : v &= z - x \end{aligned} \tag{3.4}$$

Parmak izi vektörlerinin, üretilmeleri esnasında NOMA baz ve kodlama vektörlerinin Gauss ağırlık vektörleri ile çarpımları sonucunda çok değişkenli Gauss dağılımına uydurulduğu ve elemanlar arasında bağımsızlığın sağlandığı

anımsandığında, (3.4)'de verilen hipotezlerin de aynı dağılım özelliklerini taşıdığı görülebilir. Bu doğrultuda düşünüldüğünde hipotezlerin gerçekleşme olasılıkları standart sapmaları sırasıyla σ_0 ve σ_1 olduğu durumda (3.5) ve (3.6)'daki gibi ifade edilebilir.

$$H_0 : p(n_A) = \prod_{a=1}^D \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left(-\frac{n_A^2(a)}{2\sigma_0^2}\right) \quad (3.5)$$

$$H_1 : p(z-x) = \prod_{a=1}^D \frac{1}{\sqrt{2\pi}\sigma_1} \exp\left(-\frac{(z(a)-x(a))^2}{2\sigma_1^2}\right) \quad (3.6)$$

Buna göre, iki hipotezin de eşit olasılıklı olduğu varsayımı altında, Öklid metriği için hata olasılığını en azlayan en iyi karar eşiği τ (3.7)'de verilen eşitlikle hesaplanır [7].

$$\tau = \sqrt{(2D) \log\left(\frac{\sigma_1}{\sigma_0}\right) \left(\frac{\sigma_0^2 \sigma_1^2}{\sigma_1^2 - \sigma_0^2}\right)} \quad (3.7)$$

Eşitlik (3.7)'den görülebileceği üzere, parmak izi eşleme için en iyi eşik değerinin hesaplanmasında her iki hipotezin yoğunluk fonksiyonlarının standart sapmaları kullanılmaktadır. Bu standart sapmalar sistemin bir eğitim kümesi üzerinde çıkardığı parmak izleri kullanılarak hesaplanır. Bunun için öncelikle hiçbir saldırıya uğramamış videolardan çıkarılan anahtar-çerçeveler için bir parmak izi veritabanı oluşturulur. Ardından da hem veritabanında bulunan videoların saldırıya uğrayarak bozulmuş versiyonları hem de bambaşka videolardan elde edilen anahtar-çerçeveler için parmak izleri hesaplanır. Veritabanında bulunan videoların parmak izleri ile saldırıya uğramış versiyonlarından elde edilen parmak izleri karşılıklı gelecek çiftler halinde birbirlerinden çıkarılarak n_A gürültü vektörlerine karşı düşen parmak izleri elde edilir. Benzer şekilde, veritabanında bulunan parmak izleri ile kendilerinden alakasız videolardan elde edilen parmak izleri birbirinden çıkarılarak $z-x$ fark vektörlerine karşı düşen parmak izleri elde edilir. Elde edilen gürültü vektörlerinin standart sapmaları hesaplanarak σ_0 , fark vektörlerinin standart sapmaları hesaplanarak da σ_1 elde edilir. Eşleme karar eşiği τ , eğitim aşamasında bulunan bu standart sapmalar hesaplanarak bulunur ve veritabanına kaydedilir. Arama-getirme uygulamasında, verilen sorgu videosundan çıkarılan parmak izleri ile veri tabanında

kayıtlı bulunan parmak izleri arasında Öklid uzaklığına bakılır. Bulunan uzaklık bu eşik değerin altında ise eşleme sağlanır, eğer sorgu parmak izi ile veritabanındaki parmak izlerinin hiçbiri arasındaki uzaklık bu değerin altında değilse sorgu videosunun orijinalinin veritabanında olmadığı kararı verilir.

Önerilen yöntem, NOMA sonucunda elde edilen baz ve kodlama matrislerinin giriş verisindeki global değişikliklerden fazla etkilenmemesi, dolayısı ile parlaklık, karışıklık, gamma değişimi ya da toplamsal gürültü eklenmesi gibi saldırılara dayanıklı olmasından yola çıkarak geliştirilmiştir. Ancak veri üzerine uygulanan öteleme, döndürme gibi geometrik bozulmalarda NOMA'nın bu haliyle yeterli olmadığı görülmüştür. Bu yüzden önerilen sisteme NOMA yönteminin içerisine bu tür bozulmaların etkisini katmayı amaçlayan DD-NOMA yöntemi [1] eklenmiştir. Buna göre sistem, gelen sorgu parmak izi veritabanında bulunan parmak izlerinden hiçbirine eşlenmediği takdirde gelen sorgu çerçevesi için DD-NOMA tabanlı eşleme yöntemini uygulayarak uygun sonucu arar. Bir sorgu videosunun veritabanında olmadığına ilişkin tam karar ancak bu yöntem de bir eşleme sonucu üretilmediyse verilir. Aksi takdirde DD-NOMA yöntemi ile bulunan eşleme sonucu kullanılır. DD-NOMA yöntemi ile video parmak izi çıkarımı ve eşleme bir sonraki bölümde anlatılmaktadır.

3.1.3 Dönüşümle Değişmeyen Negatif Olmayan Matris Ayırıştırma

Günlük hayatta karşılaşılan sorunlar göz önünde bulundurulduğunda NOMA yönteminin veriyi temsil özelliğinin veriyi etkilemesi olası geometrik bozulmalar göz önünde bulundurularak geliştirilmesi gerekebilir. Özellikle görüntü veya bu tezin konusunu oluşturan video kopya sezme uygulamaların girişinde, Şekil 3.2'de de örnekleri görülebilen, öteleme, döndürme vb. geometrik bozulmalara uğramış verilerle karşılaşmak çok olasıdır. İşte bu gibi durumlarda bir üstteki bölümde bahsedilen anahtar-çerçeve tabanlı parmak izi yönteminin eşleme yapamaması durumunda sisteme gelen verinin veritabanında gerçekten olmadığına karar verebilmek için bu bölüm altında anlatılan, dönüşümle değişmeyen negatif olmayan matris ayırıştırma “DD-NOMA” (*Transform Invariant NMF*, *T-NMF*) yöntemi kullanılabilir.



Şekil 3.2: (a) Orjinal bir video çerçevesi ve örnek geometrik ataklı çerçeveler (b) yatayda aynalama (c) düşeyde aynalama (d) aşağı öteleme (e) sola öteleme (f) sağa öteleme (g) yukarı öteleme (h) resim içinde resim (iki boyutlu yeniden boyutlandırma) (i) tek boyutta yeniden boyutlandırma (j) döndürme

DD-NOMA yaklaşımı özellikle verilerin bu tür bozulmalara karşı daha düzgün temsil edilmesinin bir yöntemi olarak önerilmiştir [29]. Buna ek olarak DD-NOMA yönteminin burada konu edilen anahtar-çerçeve tabanlı video parmak izi çıkarma problemlerine uygulanması da yakın zamanda önerilmiştir [1]. Böylelikle geleneksel NOMA yöntemine göre geometrik ataklara karşı gürbüzlüğün artması hedeflenmiş ve başarılı olduğu gözlemlenmiştir. Bunun için geleneksel NOMA yönteminin maliyet fonksiyonu ve güncelleme adımları, söz edilen bozulmaları temsil eden dönüşüm matrislerinin de eklenmesi ile verideki bu tür bozulmalara dayanıklı olacak şekilde güncellenmiştir.

Güncellemelerden sonra Denklem (2.2)'deki maliyet fonksiyonu aşağıdaki şekilde ifade edilir.

$$F = \|V - E\|^2 = \sum_{i=1}^p \sum_{j=1}^q \left(V_{ij} - \sum_{m=1}^M \sum_{l=1}^r (T^z W)_{il} H_{lj}^z \right)^2 + \sum_{l,j,m} H_{lj}^m \quad (3.8)$$

$$E_{ij} = \sum_{m=1}^M \sum_{l=1}^r (T^m W)_{il} H_{lj}^m \quad (3.9)$$

Baz ve kodlama matrisleri olan W ve H matrislerine ilişkin güncelleme adımları da sırasıyla (2.6) ve (2.10)'daki hallerinden aşağıdaki hale getirilir;

$$H_{ij}^m(t+1) \leftarrow H_{ij}^m(t) \frac{\sum_{i=1}^p (T^m W)_{il} V_{ij}}{\sum_{i=1}^p (T^m W)_{il} E_{ij}} \quad (3.10)$$

$$W_{il}(t+1) \leftarrow W_{il}(t) \frac{\sum_{m=1}^M \sum_{j=1}^q \left((T^m)^T \right)_{ij} H_{jl}^m}{\sum_{m=1}^M \sum_{j=1}^q \left((T^m)^T E \right)_{ij} H_{jl}^m} \quad (3.11)$$

Burada T^m , $m = 1, \dots, M$ dönüşüm matrislerini gösterirken M ; dayanıklı olunması istenen ve hesaba katılan dönüşümlerin sayısı olarak ifade edilmiştir. Ayrıca DD-NOMA bu probleme uygulandığında doğabilecek olan tutarsız çözümlerden kaçınmak ve elde edilecek kodlama matrislerini seyrekliğe zorlamak için [1]'de, daha önceden [29]'da verilmiş olan maliyet fonksiyonuna bir terim eklenerek (3.8)'deki maliyet fonksiyonunun toplam terimi elde edilmiştir.

Burada unutulmaması gereken nokta DD-NOMA yönteminin başarısının dönüşüm matrislerinin yapısı ve fonksiyonelliğine bağlı olduğudur. Söz konusu dönüşüm matrislerinin önemli özelliklerinden biri dikgen (*orthogonal*) seçilmeleridir. Böylelikle dönüşüm matrisinin temsil ettiği bozulmanın geometrik anlamda tam tersine karşı gelen bir bozulma, basitçe dönüşüm matrisinin transpozesi kullanılarak bulunabilir. Bu da DD-NOMA yönteminin (3.10) ve (3.11) güncelleme adımlarında görülen V_i ve R_i matrislerinin nasıl geometrik olarak ters bir dönüşüme uğradığını açıklar. Şekil 3.3'te aynalama (*flip*) ve öteleme (*shift*) işlemlerine karşı düşen örnek dönüşüm matrisleri görülebilir.

[1]'de önerilen parmak izi yöntemi, bir videodan alınan bir anahtar-çerçevenin DD-NOMA giriş matrisi olan V 'yi elde ederken, $(w \cdot h)$ boyutlu çerçevenin kendisi ve hesaba katılan her saldırı için geometrik olarak dönüştürülmüş versiyonlarını kullanır. Burada çerçevenin kendisi ve dönüştürülmüş versiyonları (3.12)'deki şekilde V içerisine sütun vektörler ($\mathbf{v}_{orj}$) ve ($\mathbf{v}_{dön_j}$, $j = 1, \dots, m$) olarak dizilir ve bu sayede V matrisi, yalnızca örnek görüntüyü değil, örnek görüntünün geometrik olarak dönüşmüşlerini de temsil eder hale gelir.

$$V = \begin{bmatrix} \mathbf{v}_{orj} & \mathbf{v}_{dön_1} & \dots & \mathbf{v}_{dön_m} \end{bmatrix}_{(w \cdot h) \times (m+1)} \quad (3.12)$$

$$\begin{aligned}
& \begin{bmatrix} \mathbf{E} & \mathbf{Z} & \cdots & \mathbf{Z} \\ \mathbf{Z} & \mathbf{E} & & \vdots \\ \vdots & & \ddots & \mathbf{Z} \\ \mathbf{Z} & \cdots & \mathbf{Z} & \mathbf{E} \end{bmatrix} & \begin{bmatrix} \mathbf{Z} & \cdots & \mathbf{Z} & \mathbf{I} \\ \vdots & & \mathbf{I} & \mathbf{Z} \\ \mathbf{Z} & & & \vdots \\ \mathbf{I} & \mathbf{Z} & \cdots & \mathbf{Z} \end{bmatrix} & \begin{bmatrix} \mathbf{I}^* & \mathbf{Z} & \cdots & \mathbf{Z} \\ \mathbf{Z} & \mathbf{I}^* & & \vdots \\ \vdots & & \ddots & \mathbf{Z} \\ \mathbf{Z} & \cdots & \mathbf{Z} & \mathbf{I}^* \end{bmatrix} & \begin{bmatrix} \mathbf{I}^{\leftarrow k} & \mathbf{Z} & \cdots & \mathbf{Z} \\ \mathbf{Z} & \mathbf{I}^{\leftarrow k} & & \vdots \\ \vdots & & \ddots & \mathbf{Z} \\ \mathbf{Z} & \cdots & \mathbf{Z} & \mathbf{I}^{\leftarrow k} \end{bmatrix} \\
& \text{(a)} & \text{(b)} & \text{(c)} & \text{(d)}
\end{aligned}$$

Şekil 3.3: Ayırıştırma kullanılan bazı dönüşüm matrislerinin (T^m) yapıları
 (a) yatayda aynalama (b) düşeyde aynalama (c) sola k sütun öteleme
 (d) sağa k sütun öteleme. $\mathbf{E}$, $\mathbf{Z}$ ve $\mathbf{I}$ matrisleri sırasıyla değişim matrisi, tam sıfır matrisi ve birim matris olup, $\mathbf{I}^{\rightarrow k}$ sütunları ok yönünde k adet ötelenmiş birim matristir.

V matrisinin hesaplanmasının ardından (3.8) ve (3.9)'da gösterildiği şekilde DD-NOMA gerçekleşir ve elde edilen baz matrisi W alınan çerçevenin içeriğini yegane temsil eden öznitelik vektörü olarak alınır ve arama-getirme esnasında kullanılmak üzere veritabanına kaydedilir. Burada önemli olan nokta, T^m dönüşüm matrislerinin etkisiyle giriş verisindeki benzer ancak geometrik anlamda bozulmuş içeriğin bozucu etkisi baz vektörlerine yansımaz. Bunun yerine benzer verilerin temel içerikleri baz vektörlerinde saklanır. Böylelikle, ayırıştırma sonucunda elde edilen baz vektörleri, yani W kullanılarak yapılan bir geri çatma işleminde V matrisi ve dolayısıyla video içeriklerinin farklı dönüşümlere uğramış versiyonları pek az bir kayıpla elde edilebilir. Bu özellik arama-getirme esnasında eşleme kararının alınmasında çok önemlidir, çünkü DD-NOMA tabanlı sistemde eşleme geri çatma işlemi ile sağlanır.

DD-NOMA'nın tanımı yukarıdaki şekilde verildikten sonra bu yöntemle göre çalışan bir sistemin eşleme işlemini nasıl gerçeklediği açıklanabilir. Sistem öncelikle, kendisine gelen anahtar-çerçeveyi daha önceki yöntemde olduğu gibi DC katsayıları kullanarak yeniden boyutlandırır. Yeniden boyutlandırılmış bu anahtar çerçevenin ilgilenilen dönüşümlere uğramış versiyonları T^m dönüşüm matrisleri kullanılarak hesaplanır. Ardından, yeniden boyutlandırılmış anahtar çerçeve ve dönüşüme uğramış versiyonları sütun vektörleri halinde Denklem (3.12)'deki gibi dizilip tek bir matriste toplanarak DD-NOMA yönteminin veri matrisi V elde edilir.

Sistem, elde edilen V matrisini kullanarak sırasıyla veritabanına önceden indeksleme sırasında kaydedilmiş olan her W matrisi için Denklem (3.8)'de verilen güncelleme adımını kullanarak H matrisini günceller. Güncellenmiş H matrisi kullanılarak Denklem (3.13)'te verilen geri çatma hatası hesaplanır [1].

$$Hata_{geri-\text{\u00e7atma}}^{x,y} = \left\| V^y - \sum_{m=1}^M T^m W^x H^{m,y} \right\|^2 \quad (3.13)$$

Burada r deęiřkeni DD-NOMA'nın ayırıştırma d\u00fczeyi (*rank*) iken x ve y sırasıyla referans ve sorgu \u00e7er\u00e7evelerinin indeksleridir.

Eřleme adımlarına devam edilecek olursa; sistem belirlenen bir \u00fcst tekrar sayısı veya hata miktarına kadar yukarıda bahsedilen g\u00fcncelleme ve geri \u00e7atma adımlarını tekrarlar. G\u00fcncelleme adımları bittikten sonra elde edilen geri \u00e7atma hatası eřleme i\u00e7in \u00f6nceden belirlenen bir eřięin altında kaldıysa sistem başarılıdır. Aksi durumda ise kullanılan W 'ya iliřkin g\u00f6r\u00fcnt\u00fc ile sisteme gelen g\u00f6r\u00fcnt\u00fcy\u00fc eřleyememiř sayılır ve veritabanındaki bir sonraki kayıt alınarak b\u00fct\u00fcn veritabanı uygun eřleme i\u00e7in taranmaya devam eder.

Bu y\u00f6ntem geometrik d\u00f6n\u00fcř\u00fcmle karřı başarılı olsa da temelde iki zayıflığı vardır. Bunlardan birincisi sistemin giriř verisine uygulanabilecek geometrik d\u00f6n\u00fcř\u00fcmlemin \u00f6nceden bilinmesinin gerekmesidir. DD-NOMA'da, sistemin hesaba kattığı, yani dayanıklı olmayı hedefledięi d\u00f6n\u00fcř\u00fcmlemleri ve onlara iliřkin d\u00f6n\u00fcř\u00fcm matrislerini \u00f6nceden bildięi varsayılır. Bu d\u00f6n\u00fcř\u00fcmlemler haricinde gelebilecek dięer geometrik d\u00f6n\u00fcř\u00fcmlemlere karřı y\u00f6ntemin g\u00fcrb\u00fczl\u00fcę\u00fc ortadan kaybolur.

DD-NOMA ile ilgili ikinci dikkat edilmesi gereken nokta, bu y\u00f6ntemde indeksleme esnasında veritabanına ger\u00e7ek anlamda parmak izlerinin eklenmemesidir. Bu y\u00f6ntemde veritabanına eklenen veri parmak izleri deęil, W baz matrisleridir. Bu da veritabanında idealde bir parmak izinin kaplayacaęı alandan daha fazlasının kaplanmasına neden olur. Ayrıca eřleme kriteri olarak \u00f6nceki y\u00f6ntemde olduęu gibi doęrudan bir metrik kullanılamaması ve her defasında iteratif g\u00fcncelleme adımları sonrasında geri \u00e7atma hatasının hesaplanması da sistemin arama-getirme esnasında parmak izi ile eřleme yapan y\u00f6ntemlere g\u00f6re daha yavaş \u00e7alıřmasına neden olur. Zaten bu nedenle bu \u00e7alıřmada, aslında [1] tarafından global d\u00f6n\u00fcř\u00fcmlemlere karřı da g\u00fcrb\u00fczl\u00fcę\u00fc g\u00f6sterilmiř olan DD-NOMA y\u00f6nteminin doęrudan kullanılması deęil, sadece B\u00f6l\u00fcm 3.1.1'de anlatılan sistemde uygun eřlemenin yapılamaması durumunda kullanılması \u00f6nerilmektedir. B\u00f6ylelikle DD-NOMA, 3.1.1'de \u00f6nerilen y\u00f6ntem ile birlikte kullanıldığında sistem \u00f6ncelikle ilk y\u00f6nteme g\u00f6re uygun eřlemeyi arar. řayet uygun eřleme bulunduęu takdirde DD-NOMA y\u00f6ntemine bařvurmaya gerek yoktur. İlk y\u00f6nteme uygun eřleme bulunamadığında ise sistem nihai kararı

vermeden önce gelen çerçevenin geometrik bir bozulmaya uğramış olma ihtimaline karşı DD-NOMA yöntemine başvurarak son bir denetim uygulanmış olur.

3.2 İçerik Tabanlı Video Parmak İzi Çıkarımı

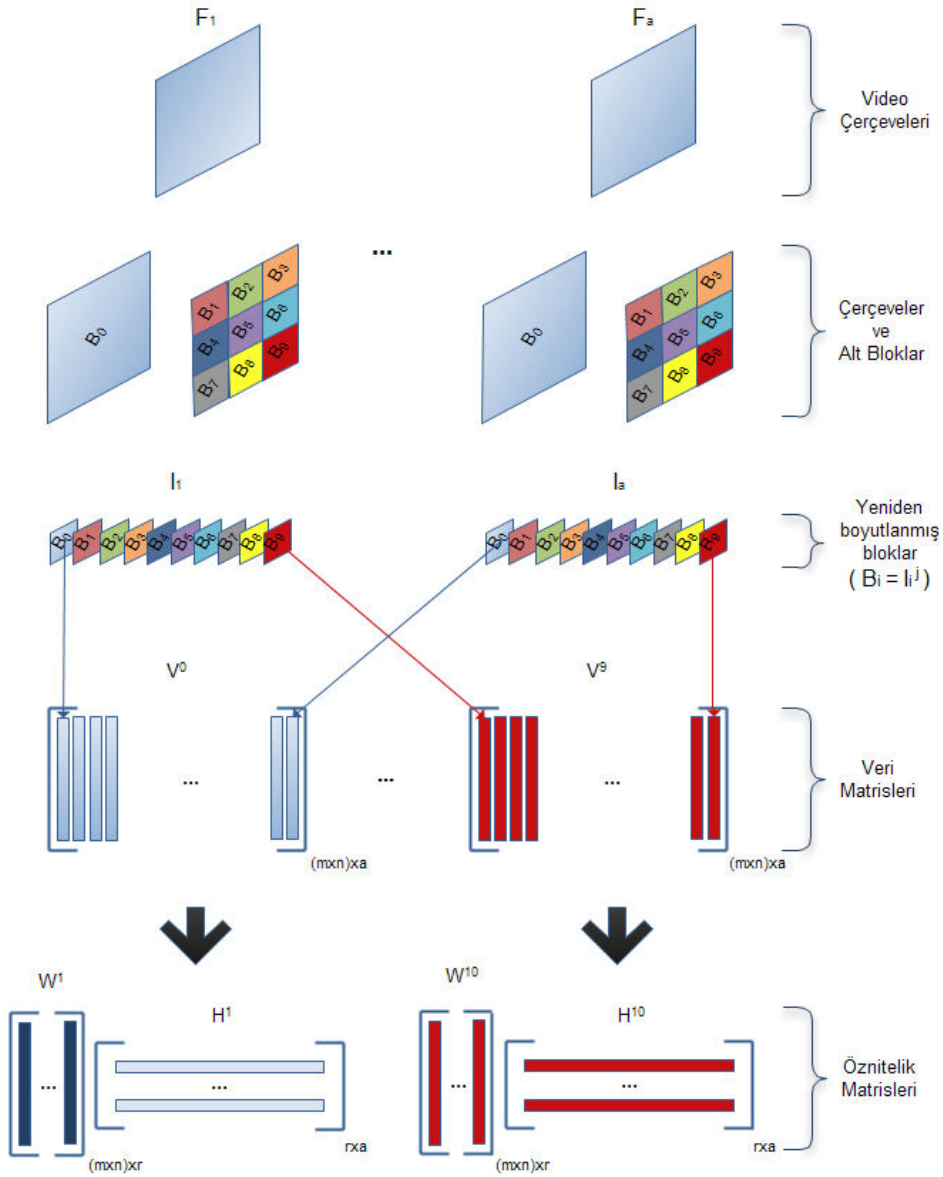
Bölüm 3.1'in başında anahtar-çerçeve tabanlı parmak izi çıkarımı anlatılırken söz konusu sistemin sadece videolardan çıkarılmış anahtar-çerçeveleri kullandığı ve bu yüzden video içeriğine rastgele erişime olanak sağladığını söylemenin mümkün olmadığı belirtilmiştir. Bu nedenle de anahtar-çerçeve tabanlı sistemlerde sorgu ve referans video klipleri arasında detaylı bir konum eşlemesi gerçekleminin de mümkün olmadığı ve bu problemi çözmek için alternatif bir video parmak izi çıkarma yöntemi önerileceği de belirtilmiştir. Bu bölümde, bahsedilen problemi çözmek için video çerçevelerini birbirlerini aralıksız takip eden gruplar halinde ele alan ve bir grup video çerçevesi kullanarak bir video klbinin alt parçalarını için parmak izi üreten bir yöntem anlatılmaktadır

Önerilen sistemde her bir parmak izi, video içerisinde önceden karar verilen minimum bir süre için çıkarılır. Bu minimum süre Bölüm 1'de bir örnekle anlatıldığı gibi parmak izi sisteminin çözünürlük (*granularity*) ölçütüdür. Sistem bu süreden daha kısa süreler için eşleme veya konumlandırma yapamaz.

Sistem, ilk aşamada çözünürlük ölçütündeki süreyi temsil edecek kadar sayıda video çerçevesini bir tampona alır. Alınan bu $F_i \in R^{w \times h} \quad i = 1, \dots, a$ çerçevelerin her biri boşluk kalmayacak şekilde 9 eş alt parçaya bölünür. Öznitelik çıkarımında çerçevenin kendisinin yanısıra bu alt parçaların da kullanımı ile video verisinin daha detaylı temsil edilmesi amaçlanmıştır. Ayrıca böylelikle video çerçevelerinin belli bölgelerini etkileyip geri kalan bölgelerine dokunmayan altyazı, kanal logosu gibi örüntü ekleme veya kırpma gibi bir takım saldırı türlerine karşı da gürbüzlüğün artırılması amaçlanmıştır. Bu alt parçalar elde edildikten sonra her bir çerçeve ve o çerçevelerden alınan parçalar anahtar-çerçeve bazlı parmak izi çıkarma yönteminde olduğu gibi DC katsayıları kullanılarak $w \times h$ 'dan $p \times q$ 'ya yeniden boyutlandırılır. Böylelikle Şekil 3.4'te görülen her bir çerçeveyi temsil eden yeniden boyutlanmış 10 adet eş boyutlu matris (blok) $I_i^j \in R^{p \times q} \quad i = 1, \dots, a \quad j = 1, \dots, 10$ elde edilmiş olur.

İkinci aşamada bu yeniden boyutlanmış bloklar ve çerçeveler kullanılarak NOMA yönteminde kullanılacak olan $V^j \in R^{(p \times q) \times a} \quad j = 1, \dots, 10$ veri matrisleri inşa edilir.

Şekil 3.4'te görülebilen bu dördüncü aşamada her blok için bir adet olmak üzere 10 adet veri matrisi oluşturulur ve veri matrislerinin sütunları da birbirini takip eden çerçevelerden çıkarılmış olan ilgili blokların sütun vektörleri olarak dizilmesi ile oluşturulur.



Şekil 3.4 : Video çerçevelerinden öznitelik matrislerinin çıkarılması

Böylelikle ilk iki adım sonucunda NOMA yönteminde kullanılmak üzere hazırlanmış 10 adet veri matrisi elde edilir. Her bir veri matrisi $V^j \in R^{p \times q}$, $j = 1, \dots, 10$ olmak üzere NOMA kullanılarak, $W^j \in R^{p \times r}$ karışım/baz matrisi ve $H^j \in R^{r \times q}$ kodlama matrisi şeklinde iki matrisin çarpımına ayrıştırılır. Böylelikle her parmak izi süresi için 10'ar W, H çifti elde edilmiş olur. Elde edilen bu çiftler Şekil 3.4'te son

aşamada gösterilmiştir. Bu W^j ve H^j matrisleri videonun karşı gelen bölümüne ilişkin öznitelikler olarak tanımlanır ve parmak izi çıkarımında bu özniteliklerden yararlanır.

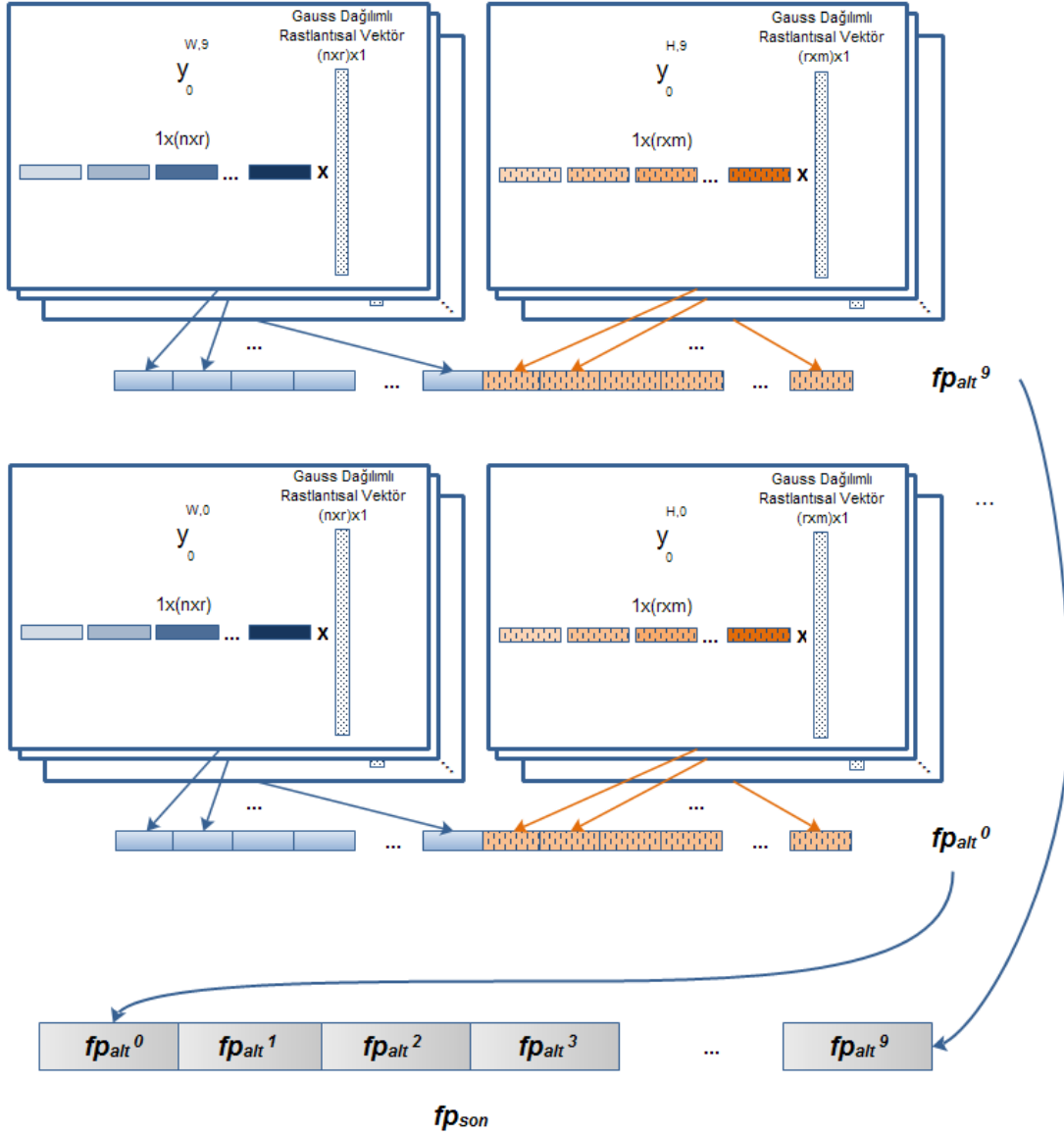
W^j ve H^j matrisleri elde edildikten sonra bu öznitelikler kullanılarak anahtar-çerçevesel yöntemdekine benzer şekilde bir parmak izi oluşturulur. Bunun için öncelikle bu matrisler Denklem (3.14)'daki şekilde satır vektörleri haline getirilir:

$$\begin{aligned} y^{w_j} &= [W_{k,1}^j \quad W_{k,2}^j \quad \dots \quad W_{k,r}^j]^T \quad k = 1, 2, \dots, m \\ y^{h_j} &= [H_{1,l}^j \quad H_{2,l}^j \quad \dots \quad H_{r,l}^j]^T \quad l = 1, 2, \dots, n \end{aligned} \quad (3.14)$$

Burada y^w ve y^h vektörleri sırasıyla $m \times r$ ve $r \times n$ boyutludur. Satır vektörleri haline gelmiş olan bu öznitelik matrisleri anahtar-çerçeve tabanlı parmak izi çıkarma yönteminde olduğu gibi sözde-rastlantısal Gauss ağırlık vektörleri ile Denklem (3.15)'de verildiği ve Şekil 3.5'te gösterildiği gibi çarpılarak ağırlıklandırılır ve $D=N+M$ boyutlu, kayan-noktalı sayı tipinden parmak izi vektörü $fp_{alt}^j \in R^D$ elde edilmiş olur.

$$fp_{alt}^j(k) = \begin{cases} \sum_{l=1}^{m.r} y_l^w g_l^w, & k = 1, 2, \dots, N \\ \sum_{l=1}^{n.r} y_l^h g_l^h, & k = (N+1), (N+2), \dots, (N+M) \end{cases} \quad (3.15)$$

Bu işlemin amacı aynı bir önceki anahtar-çerçeve bazlı yöntemde olduğu gibi hem parmak izinin elemanları birbirinden bağımsız Gauss dağılımına uydurulmasını sağlayıp eşleme karar eşiği bulunmasını kolaylaştırmak hem de ciddi bir boyut düşürme performansına erişmektir. Böylelikle özellikle sistemin pratikte hem daha hızlı çalışması hem de parmak izi veritabanı boyutunun düşük olması amaçlanmaktadır. Buradan görüleceği üzere önerilen video parmak izi çıkarma yöntemi önemli bir boyut düşürme performansı sağlamaktadır. Başlangıçta alınan a adet $w \times h$ boyutlu video çerçevesi için sonuçta kayan noktalı sayı tipinden D boyutlu tek bir parmak izi vektörü elde edilir.



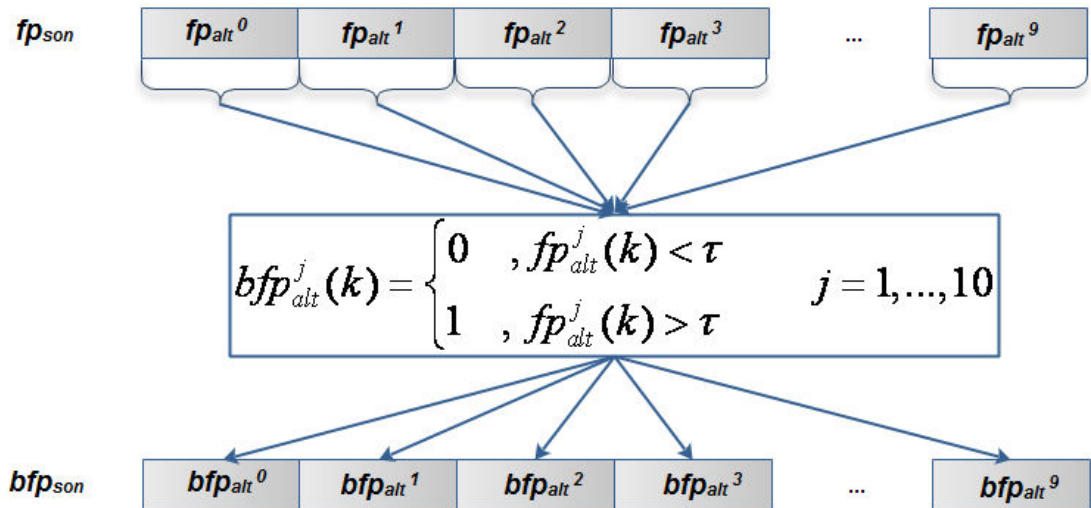
Şekil 3.5 : Satır vektörlerinin ağırlıklandırılması

Elde edilen kayan noktalı sayı tipinden parmak izlerinin yanısıra, bu çalışmada ayrıca ikili bit dizilerinden oluşan bir video parmakizi kullanımı da önerilmektedir. Bunun nedeni gerçek sistemlerde yaşanan gerek veritabanı boyutu gerekse arama-getirme süresi uzunluğu kısıtları nedeniyle kayan noktalı sayılardan oluşan parmak izlerden hem daha az yer kaplayacak hem de daha hızlı çalışacak temsiller elde etmektir. Bu ikili parmak izlerinin elde edilmesinde [30] tarafından önerilen sayısal damgalama yönteminde kullanılmış ve [5] tarafından video parmak izi çıkarımında da kullanılabileceği gösterilmiş olan ikilileştirme yöntemi kullanılmıştır. Bu yöntemle göre, çıkarılan parmak izinin taşıdığı bilgi miktarının en yüksek olması amacıyla parmak izi bitlerinin yarısının sıfırlar kalan yarısının da birlerden oluşması sağlanmalıdır. Bir diğer ifade ile parmak izi vektöründeki sıfırlar ve birlerin sayısı

eşit olmalıdır. Bunu sağlamak için de vektördeki ortanca değere sahip sayı eşik değeri seçilerek vektördeki değerler bu sayıya göre eşiklenir. Bu τ eşiklerinin üzerinde kalan bitlere 1, eşik altında kalan sayılara 0 değeri atanır. Bu çalışmada elde edilmiş olan video parmak izlerinin ikilileştirilmesinde de bu yöntem önerilmektedir. Ancak burada dikkat edilmesi gereken nokta, bu yöntemin video parmak izinin tamamı üzerinde değil, Eşitlik (3.16) ve Şekil 3.6’da görüldüğü gibi alt parçalar halinde her blok parmak izi fp_{alt}^j için ayrı ayrı yapılması ve ardından (3.17)’de görüldüğü gibi bu alt parmak izlerinin birleştirilmesinden bütün parmak izi bfp ’nin elde edilmesidir.

$$bfp_{alt}^j(k) = \begin{cases} 0 & , fp_{alt}^j(k) < \tau \\ 1 & , fp_{alt}^j(k) > \tau \end{cases} \quad j = 1, \dots, 10 \quad (3.16)$$

$$bfp = [bfp_{alt}^1 \quad bfp_{alt}^2 \quad \dots \quad bfp_{alt}^{10}] \quad (3.17)$$



Şekil 3.6 : Ortanca değere göre eşikleme ile ikili parmak izinin elde edilmesi

Bu yöntem ile elde edilecek video parmakizleri orijinal videoyu yegane bir şekilde indekslerken veritabanı boyutu da kayan noktalı sayılı parmak izlerinden oluşan bir veritabanına göre çok daha düşük olur. Bunların yanısıra önerilen bu ikili sistem kayan noktalı sayılı sisteme göre daha hızlı arama-getirme avantajına da sahiptir.

4. TESTLER

Bu bölümde, Bölüm 3’de detaylı olarak anlatılan ve negatif olmayan matris ayrıştırma (NOMA) yöntemi temel alınarak oluşturulan kopya sezme sistemlerinin başarımı irdelenmektedir. Deneysel sonuçlara geçilmeden önce test kümelerinin seçimi ve sonuçların raporlanmasında kullanılan başarımların kriterlerine değinilecek ardından sırasıyla Bölüm 3.1’de bahsedilen anahtar-çerçeve tabanlı parmak izi sistemi ve Bölüm 3.2’de anlatılan video parmak izi sistemlerine ilişkin test sonuçları verilecektir.

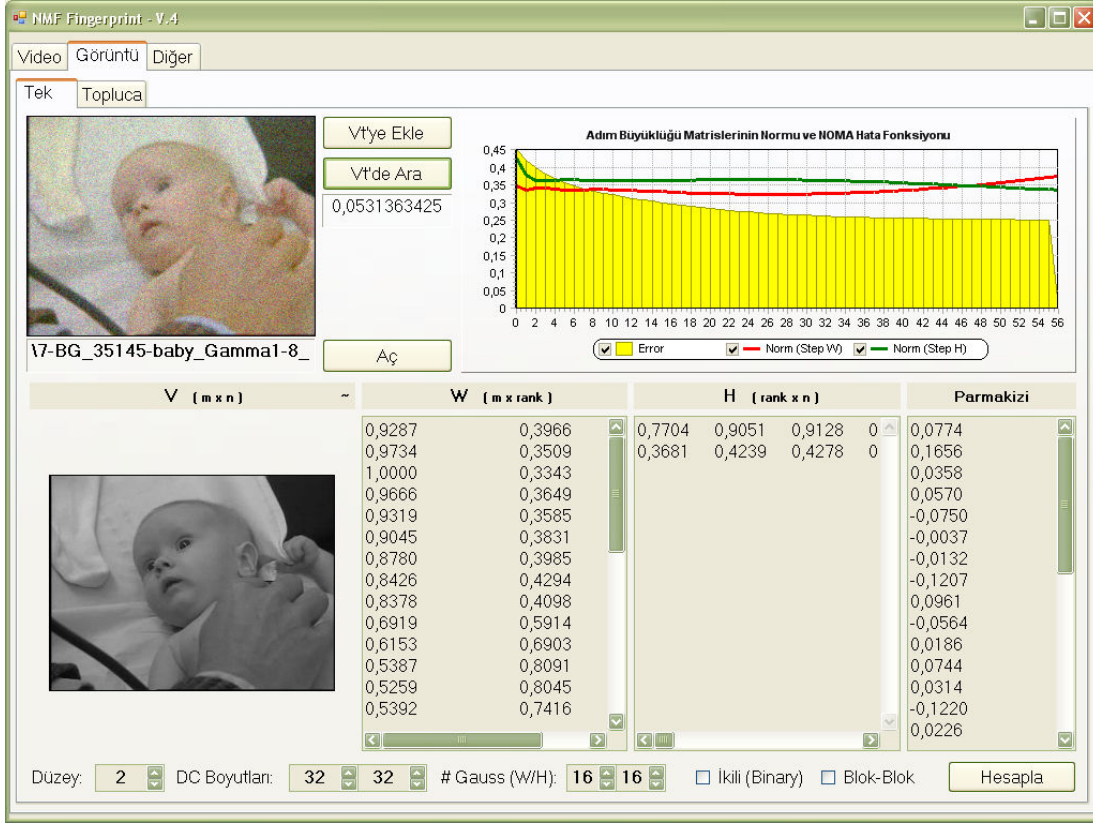
Testler, görüntü ve video işleme algoritmaları C/C++ dillerinde, kullanıcı arayüzü ise C# dilinde hazırlanan bir yazılım ile denenmiştir. Yazılımda MPEG biçimindeki video dosyalarının açılabilmesi için FFmpeg [31], görüntü işleme arkaplan fonksiyonlarının gerçekleştirilmesi amacıyla da OpenCV [32] yazılım kitaplıkları kullanılmıştır. Testlerin gerçekleştirildiği parmak izi veritabanı ise MSSQL Server 2000 ortamında hazırlanmıştır. Yazılım, video kliplerini olduğu gibi ya da gerektiğinde video çerçevelerini ayrı ayrı seçmeye olanak veren, NOMA parametrelerinin ve parmak izi uzunluğunun ayarlanabildiği, oluşturulan parmak izlerini ilgili veritabanına kayıt edip, o veritabanı üzerinde sorgulama yapabilen bir yapıdadır. Yazılımın kullanıcı arayüzü Şekil 4.1’de verilmiştir.

4.1 Başarım Kriterleri

Bu çalışma kapsamında önerilen video parmak izi çıkarma sisteminin başarımı TRECVID uluslararası ‘*benchmarking*’ grubunun [2] kopya sezme (*copy detection*) alt grubu çalışmaları için belirlenen başarımların kriterlerine uygun olarak değerlendirilmiştir. Bu kriterler aşağıda özetlenmektedir.

TRECVID tarafından her yıl katılımcılara toplam uzunluğu ve adedi değişen video klipleri test amaçlı olarak sağlanır. Sağlanan video kliplerinin bir kısmı veritabanı oluşturma amaçlı referans videolar iken, bir kısmı da arama-getirme işlemlerinde kullanılacak sorgu videolarıdır. Katılımcılara sağlanan bu video veritabanında temel olarak 3 sorgu videosu tipi vardır:

- Veri tabanında yer alan referans video kliplerinden alınan parçalara atak uygulayarak oluşturulan sorgu videoları
- Veri tabanında olmayan video kliplerinden alınmış sorgu videoları
- Veri tabanında olmayan bir video klibinin başına, sonuna ya da arasına referans video kliplerinden birinin ataklı bir parçası eklenerek hazırlanmış sorgu videoları



Şekil 4.1: Yazılımın kullanıcı arayüzü. Sol üstte bozulmaya uğramış sorgu videosunun bir çerçevesi görünürken sol altta veritabanından eşlenen referans videodaki ilgili çerçeve görünmektedir.

Sorgu videolarına uygulanan ve video parmak izlerinin dayanıklılık testlerinde kullanılan global ve geometrik saldırılara örnekler Şekil 4.2’de görülmektedir. Şekil 4.2.a’da yer alan orjinal video çerçevesine, piksel parlaklık değerlerinde bozulmaya neden olan kontrast değişimi, bulanıklaştırma, yeniden kodlama, gürültü ekleme ve gama bozulması gibi saldırıların ayrı ayrı uygulanmasından sonra elde edilen örnek video çerçeveleri Şekil 4.2.b’den Şekil 4.2.f’ye; geometrik bozulmalara neden olan resim-içi-resim, örüntü ekleme, aynalama (*flip*), kaydırma (*shift*), kırpma (*crop*) ve yeniden boyutlandırma (*resize*) saldırıları ise Şekil 4.2.g’den Şekil 4.2.l’ye gösterilmiştir.



Şekil 4.2: (a) Orjinal video çerçevesi ve TRECVID testlerinde kullanılan ataklar (b) kontrast değişimi (c) bulanıklaştırma (d) yeniden kodlama (e) gürültü ekleme (f) gama bozulması (g) resim-içi-resim (h) örüntü ekleme (i) aynalama (j) kaydırma (k) kırpma (l) yeniden boyutlandırma

TRECVID tarafından sağlanan sorgu videoları büyük oranda bu bozulmaların birden fazlasının birlikte kullanılması ile oluşturulmuştur. Bu çalışma kapsamında önerilen yöntemlerin başarımının her bir saldırı türü için daha iyi ölçülebilmesi adına, testlerde doğrudan TRECVID’den alınan sorgu videoları değil, TRECVID’den alınan referans videoları üzerine TRECVID’de belirtilen saldırı türlerinin tek tek bozulmalar halinde uygulanması ile elde edilen sorgu videoları kullanılmıştır. Böylelikle önerilen sistemlerin her bir saldırı türüne karşı dayanıklılığının ayrı ayrı gözlemlenebilmesi amaçlanmıştır.

Testlerde başarımlar temelde üç istatistikî bilgi ile raporlanmıştır. Bunlardan ilki literatürde hatırlama (*recall*) diye geçen ve doğru olarak eşlenmesi gereken referans-sorgu çiftlerinin ne oradan eşlenebildiğinin ölçütüdür. Testlerde kullanılan ikinci ölçüt doğruluk/isabetlilik (*precision*) diye geçer ve sistem tarafından toplam eşlenen referans-sorgu çiftlerinin ne kadarının gerçekten eşlenmesi gerektiğini gösterir. Bu iki ölçüt birbirlerine doğrudan bağlı olmamakla birlikte aralarında bir ters orantı olduğu söylenebilir. Test sonuçlarında bu iki ölçütün birden olabildiğince yüksek çıkması hedeflenir.

Testlerde kullanılan üçüncü başarımlı ölçütü hatırlama ve doğruluk ölçütlerinin harmonik ortalaması olan normalize kopya tespit maliyet oranı (normalized detection cost rate, NDCR) olarak seçilmiştir. NDCR aynı zamanda TRECVID’de resmi sonuçları raporlanırken kullanılan ölçüttür. NDCR, 0 ile 1 arasında değişen bir ölçüt olup ideal bir sistemde değerinin 0, tam tersi olup sürekli hatalı karar veren bir sistemde ise değerinin 1 olması beklenir. NDCR’ın formülü Denklem (4.1)’de verilmiştir.

$$NDCR = P_{Kaçırma} + \frac{C_{YanlışAlarm}}{C_{Kaçırma} \cdot R_{Hedef}} \cdot R_{YanlışAlarm}$$

$$P_{Kaçırma} = FN / N_{Hedef}$$

$$R_{YanlışAlarm} = FP / T_{Sorgular}$$
(4.1)

Denklem (4.1)’de görülen $C_{Kaçırma}$ ve $C_{YanlışAlarm}$ sırasıyla ıskalama (kaçırma) ve hatalı eşleme maliyetleridir ve TRECVID’in 2009 standartlarına göre sırasıyla 10 ve 1 değerlerini alırlar. $P_{Kaçırma}$ ve $R_{YanlışAlarm}$ değerleri ise koşullu kaçırma olasılığı ve hatalı eşleme oranına karşı düşer. Bu iki değer hesaplanmasında kullanılan FN referans video içerdiği halde veri tabanındaki referans video ile eşlenemeyen sorgu videosu sayısını, FP ise içermediği bir referans videosu ile eşlenen sorgu videosu sayısını vermektedir. N_{Hedef} , veri tabanında yer alan bir video klibinden bir parça içeren sorgu videosu sayısıdır. $T_{Sorgular}$ ise tüm sorgu videolarının toplam uzunluğudur. R_{Hedef} yine TRECVID 2009 standartlarına göre 20 değerini alır ve gözlemlerin karşılaşımla oranını yani hedef önbilgisini ifade eden bir sabittir.

4.2 Anahtar-Çerçeve Tabanlı Parmak izi Sistemi için Test Sonuçları

Bu bölümde Bölüm 3.1’de anlatılan anahtar-çerçeve tabanlı parmak izi sistemine ilişkin test sonuçları verilmiştir. Bu yöntem aşağıda raporlanan sonuçlar ile birlikte IEEE Uluslararası Görüntü İşleme Konferansı’nda (*IEEE International Conference on Image Processing, ICIP 2010*) yayımlanmıştır [33].

Önerilen sistemin performansı TRECVID video veritabanından alınan videolardan çıkarılan 211 anahtar çerçeve için 993 sorgu çerçevesi üzerinde incelenmiştir. Önerilen yöntem renk bilgisini kullanmadığından çerçeveler gri düzey olarak

kullanılmıştır. Kullanılan video çerçeveleri 32x32 piksellik DC görüntüler haline dönüştürülmüş ve NOMA düzeyi (*rank*) 2 olarak alınmıştır. Elde edilen parmak izleri kayan noktalı sayı tipinden olup eşlemede kullanılacak olan metrik Bölüm 3.1’de belirtildiği gibi Öklid uzaklığıdır. Öklid metriği için eşik değeri Bölüm 3.1.1’de anlatılan ikili hipotez testi kullanılarak elde edilmiş olup değeri 0,1293 olarak belirlenmiştir. Şayet uzaklık bu değer üzerinde çıkarsa sistem DD-NOMA yöntemine başvurur. DD-NOMA yönteminde eşlemede kullanılan geri çatma hatası için [1][6]’da önerilmiş bir değer olmamakla birlikte testler sonucunda gözlenen değerler doğrultusunda bu çalışmada geri çatma hatası için eşik değeri 0,5 olarak seçilmiştir. Parmak izi kullanılarak eşlenememiş bir çerçeve için geri çatma hatası da bu eşiğin üzerinde çıkarsa çerçevenin veritabanında olmadığına karar verilir.

Öncelikle değerlendirmeler esnasında açıklayıcı olması açısından Çizelge 4.1’de bir anahtar çerçeveye farklı ataklar uygulanarak elde edilmiş sorgu çerçeveleri ile yapılmış sorgu sonuçları verilmiştir. Bu çizelgede soldaki iki sütun sorgu videolarına uygulanan saldırı türlerini vermektedir. Ardından gelen ‘Yöntem’ sütununda ise veritabanında aramanın hangi yöntemle yapıldığı görülür. ‘Eşleme Uzaklığı’ olarak adlandırılan sütunda ise sorgu videosundan çıkarılan parmak izlerinin, eşlenmeleri gereken referans parmak izi ile olan öklid uzaklıkları yer alır. ‘Ortalama Uzaklık’ olarak adlandırılan sütunda ise sorgu videosundan çıkarılan parmak izlerinin, eşlenmeleri gereken parmak izi haricinde kalan diğer parmak izleri ile olan ortalama uzaklıkları raporlanmıştır. Bir parmak izi sisteminin düzgün çalışabilmesi için bu uzaklıkların eşleme uzaklığından daha büyük olması gerekmektedir. Çizelgede eşleme öklid uzaklığının eşik değeri olan 0,1293’ten az olduğu hücreler kırmızı/kalın olarak işaretlenmiştir. Sorgu çerçevelerinin bu yöntemle eşleşemediği durumlarda ilgili çerçeve için DD-NOMA yöntemi uygulandığından çizelgede ilgili sorgu satırlarına DD-NOMA ile geri çatma hataları da eklenmiştir. Geri çatma hatasının DD-NOMA için seçilen 0,5 eşik değerinden düşük olduğu yerlerde eşleme yapılır ve çizelgede mavi/kalın olarak işaretlenmişlerdir.

Çizelge 4.1 incelendiğinde global ataklara (S1-S6) maruz kalmış sorgu çerçevesinin NOMA tabanlı parmak izi çıkarma yöntemi kullanılarak büyük bir doğruluk ile eşlendiği görülebilir. Ancak yöntemin özellikle ciddi geometrik ataklara karşı dayanıklı olduğunu söylemek mümkün değildir. Bu sıkıntıyı gidermek için NOMA yöntemiyle eşleme yapılamayan yerlerde DD-NOMA kullanılması önerilmiştir.

Çizelge 4.1 : Anahtar-çerçeve tabanlı parmak izi yöntemine ilişkin eşleme uzaklıkları

	Saldırı Türü	Yöntem	Eşleme Uzaklığı	Ortalama Uzaklık
S1	Original	NOMA	0	0,24
S2	Gauss Bulanıklığı 3x3	NOMA	0,003	0,24
S3	Gamma 0,3	NOMA	0,164	0,26
		DD-NOMA	0,14	
S4	Gamma 0,8	NOMA	0,03	0,23
S5	Gamma 1,8	NOMA	0,08	0,24
S6	Gauss Gürültüsü (0,2)	NOMA	0,02	0,23
S7	Sola Öteleme 25 px	NOMA	0,12	0,26
S8	Sola Öteleme 50 px	NOMA	0,23	0,25
		DD-NOMA	0,04	
S9	Aşağı Öteleme 40 px	NOMA	0,1901	0,25
		DD-NOMA	0,05	
S10	Yeniden Boyutlandırma %75	NOMA	0,11	0,24
S11	Sola Döndürme 10°	NOMA	0,14	0,26
		DD-NOMA	0,67	
S12	Sola Döndürme 30°	NOMA	0,24	0,27
		DD-NOMA	0,05	
S13	S5 + S2 + S6	NOMA	0,08	0,24
S14	S3 + S2 + S6	NOMA	0,147	0,24
		DD-NOMA	0,01	

Çizelge incelendiğinde DD-NOMA yönteminin geometrik saldırılara karşı önemli ölçüde dayanıklı olduğu görülebilir. Böylelikle her iki yöntemin bir arada kullanılmasından ortaya çıkan sistemin hem global hem de geometrik saldırılara karşı dayanıklı olduğu görülebilir. Bu yöntemin zayıf yönü ise Bölüm 3.1.3’de de bahsedildiği gibi, DD-NOMA yönteminin karşılaşılabileceği saldırı türlerine ilişkin dönüşüm matrislerini önceden bilmesinin gerekmesidir. Çizelgede (S12)’de görülen döndürme işleminin başarıyla eşlenirken (S11)’de görülen ve kendisinden daha küçük bir açı değerine sahip olan döndürme saldırısının eşlenememesinin nedeni budur. Testlerde kullanılan sistemde 30° için dönüşüm matrisi tanımlıyken 10° için bir dönüşüm matrisi tanımlı değildir ve aradaki bu fark da saldırının DD-NOMA ile eşlenebilmesinin önüne geçmiştir.

NOMA tabanlı parmak izi yöntemi için yapılan bir diğer test de sistemi çeşitli global saldırılar altında sınavı eşleme başarısını ölçmektir. Buna göre sistem sırasıyla bütün global atakların bir arada bulunduğu bir test kümesi ve her bir saldırı türü için ayrı ayrı hazırlanmış test kümeleri ile denenmiş, sistemin doğruluk ve hatırlama

yüzdeleri raporlanmıştır. Veritabanı yine 211 anahtar çerçeveden oluşurken toplam sorgu çerçevesi sayısı için 993'tür. Buna göre sistemin başarımı Çizelge 4.2'deki gibi olmuştur.

Çizelge 4.2 : Anahtar-çerçeve tabanlı parmak izi sisteminin global saldırılara karşı başarımı

Saldırı Türü	Doğruluk	Hatırlama
Bütün Globaller	% 94,43	% 87,97
Gauss Bulanıklaştırma	% 100	% 99,7
Gürültü Ekleme	% 100	% 100
Gama Değişimi (1)	% 95,06	% 72,02
Gama Değişimi (2)	% 99,69	% 95,81
Parlaklık (1)	% 89	% 79,94
Parlaklık (2)	% 95,21	% 89,22
Karşıtlık	% 96,6	% 85,03

Çizelge 4.2'de gösterilen saldırılar için parametreler sırasıyla verilecek olursa; Gauss bulanıklaştırma için kullanılan ızgara boyutları 3x3 ile 11x11 arasında, gürültü ekleme için eklenen Gauss gürültüsünün standart sapması 0,1 ile 0,2 arasında, gama değişiminin (1) ile gösterilen ilk kümesi için gama değerleri 0,3 ile 1,8, (2) ile gösterilen ikinci kümesi için 0,8 ile 1,2 arasında, parlaklık değişiminin (1) ile gösterilen ilk kümesi için parlaklık değişimi ± 50 gri düzeyi, (2) ile gösterilen ikinci kümesi için ± 38 gri düzeyi arasında ve karşıtlık için $\pm \% 30$ arasında seçilmiştir. Burada gözlemlenen sonuçlar bir önceki çizelgedeki sonuçları doğrular nitelikte olup, önerilen NOMA tabanlı parmak izlerinin global saldırılara karşı etkinliğini göstermektedir.

4.3 İçerik Tabanlı Video Parmak izi Sistemi için Test Sonuçları

Bu bölümde Bölüm 3.2'de önerilen video parmak izi sisteminin başarımı irdelenmiştir.

Önerilen sistem TRECVID 2010 video veritabanından alınan toplam süreleri 11 dakika 5 saniye olan 7 adet video klibi ve bu kliplerin 11'er adet global, 15'er adet de geometrik saldırıya uğramış versiyonları kullanılarak denenmiştir. Bir diğer ifade ile testlerde 7 referans, 182 de sorgu videosu kullanılmıştır. Testlerde parmak izleri 1'er

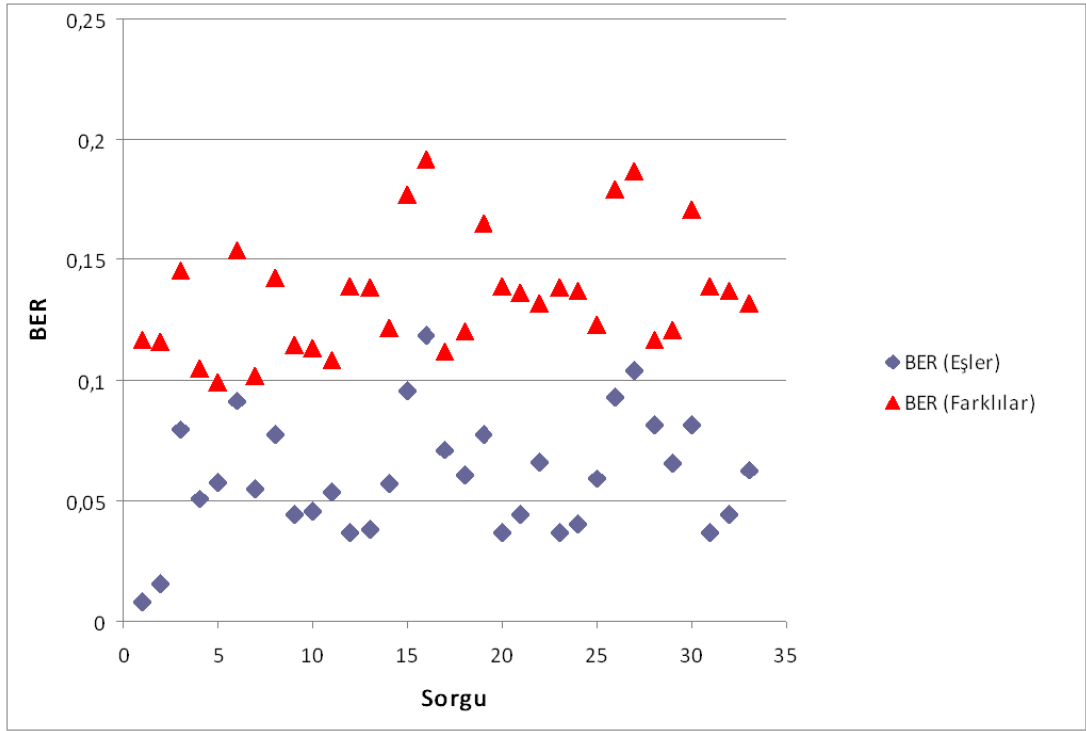
saniyelik, ardışıl ve birbirleriyle örtüşmeyen video parçaları için çıkarılmıştır. Böylelikle her videonun tamamı kapsanmış ve her bir video saniye cinsinden uzunluğu kadar alt parmak izi ile temsil edilmiştir. Parmak izlerinin 1'er saniyelik ardışıl ve örtüşmeyen parçalardan çıkarıldığı bu durumda referans parmak izi sayısı 665, sorgularda kullanılan parmak izi sayısı da global saldırılar için 7315, geometrik saldırılar için 9975 olmuştur.

Sistemde parmak izleri çıkarılırken video çerçeveleri ve bu çerçevelerden Bölüm 3.2'de anlatıldığı şekilde alınan 9 alt blok 32x32 piksellik DC görüntüler halinde kullanılmıştır. Denklem (3.17)'de görülen N ve M parametreleri eşit ve 16 olarak alınmıştır. Böylelikle her bir bloğa karşı düşen alt parmak izi 32 bit, 1 saniyelik parçayı temsil eden nihai parmak izi ise 320 bit olarak elde edilmiştir.

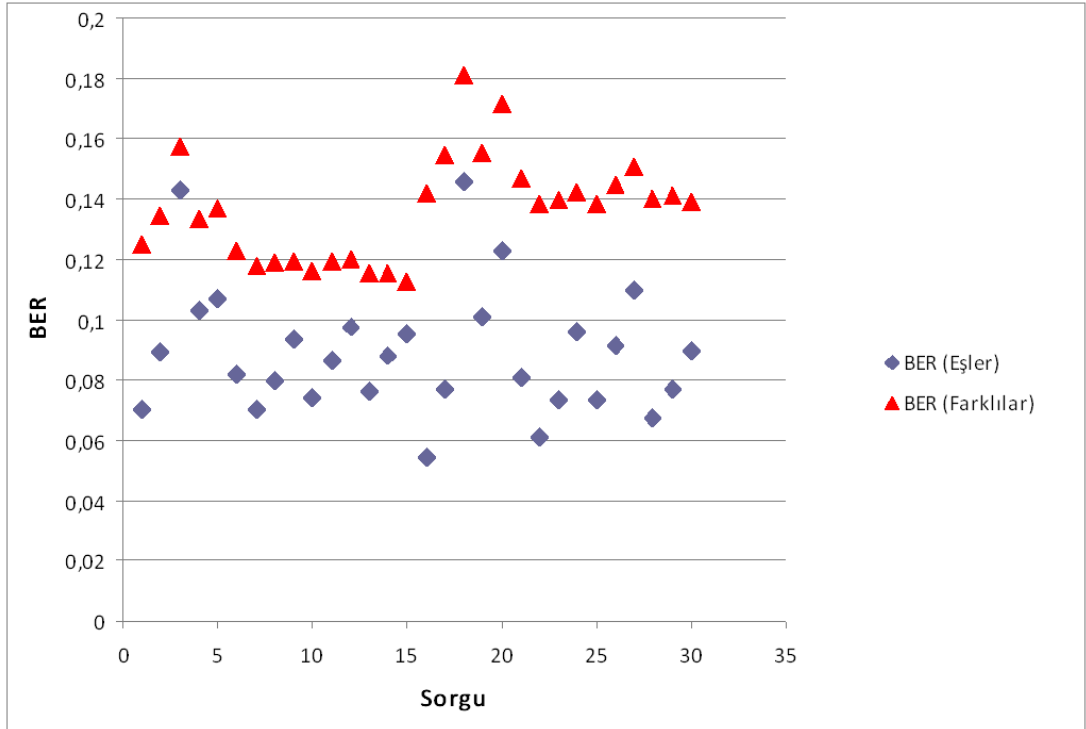
Sistemin performansı Bölüm 4.1'de belirtildiği üzere doğruluk, hatırlama ve NDCR kriterlerine göre ölçülmüş, ayrıca global ve geometrik bozulmalara karşı bit hata oranları (*bit error rate- BER*) verilmiştir. Sistemin bit hata oranı hesaplanırken iki farklı durum ele alınmıştır. Bunlardan ilki gerçek anlamda bit hata oranı olup, orjinal bir video parçasından çıkarılmış parmak izi ile onun saldırıya maruz kalmış bir parçasından çıkarılan parmak izi arasındaki farkın ölçütüdür ve sonuçlarda BER (Eşler) olarak verilmiştir. İkinci olarak verilen BER (Farklılar) ise birbiri ile alakası olmayan parmak izleri arasındaki farkın ölçütüdür (bkz. Denklem (4.2)).

$$BER = \frac{\text{HammingUzaklığı}(fp_{\text{orjinal}}, fp_{\text{bozulmuş/farklı}})}{\text{Parmak izi uzunluğu}} \quad (4.2)$$

Bir parmak izi sisteminin düzgün çalışabilmesi için BER (Farklılar)'ın BER (Eşler)'e göre daha yüksek çıkması ve bu iki sınıfın birbirinden kolaylıkla ayrılabilmesi gereklidir. Bu durumda, sistemin ürettiği parmak izlerinin saldırıya uğramış da olsa birbirinin eşi olan videolar için diğer videolar için üretilen parmak izlerine kıyasla birbirlerine daha yakın olduğu söylenebilir. Şekil 4.3 ve Şekil 4.4'te sistemin global ve geometrik ataklı durumlardaki bit hata oranları görülebilir. Şekil 4.5'te ise bu iki grafiğin birleşimi verilmiştir. Grafiklerde yatay eksen sorgu sayısı iken, dikey eksen bit hata oranı (BER) olarak verilmiştir.



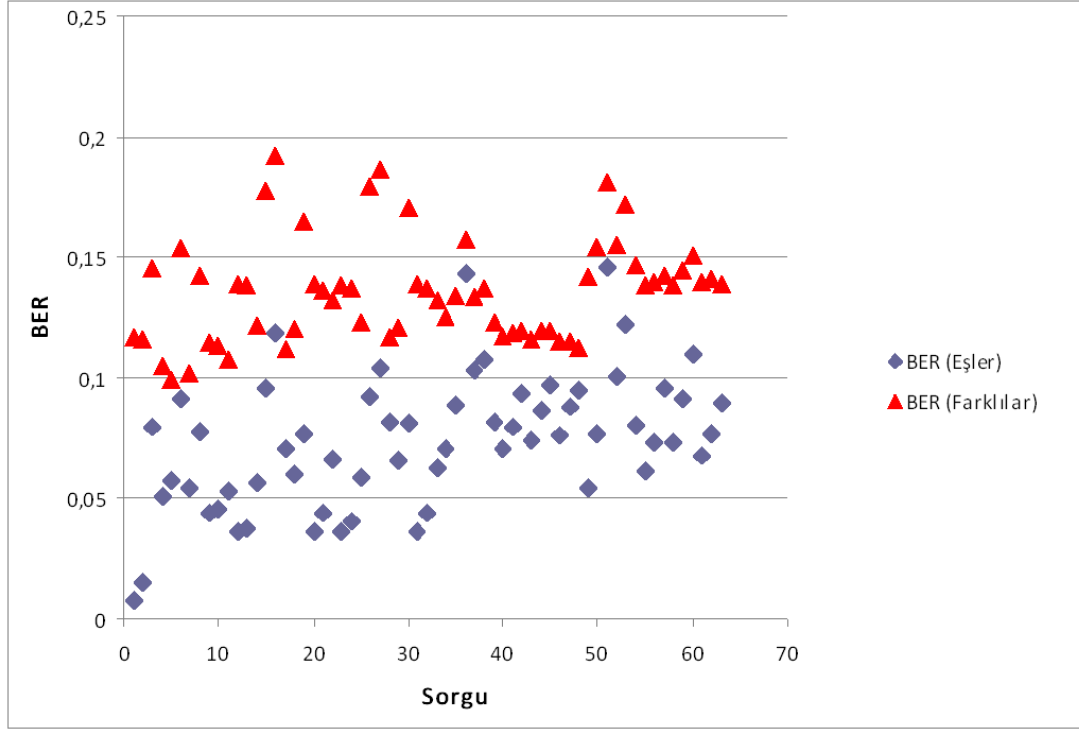
Şekil 4.3 : Global saldırılar için BER



Şekil 4.4 : Geometrik saldırılar için BER

Şekil 4.3'ten görülebileceği gibi sistemin global saldırılara karşı ürettiği bit hata oranı, birbirleri ile ilgisiz video parçalarına üretilen parmak izlerinin birbirlerinden farkından daha düşüktür ve bu iki farklı durum birbirlerinden büyük bir doğrulukla

ayrılabilir. Bu da sistemin global saldırılara karşı büyük oranda başarılı olmasını sağlar. Şekil 4.4'e bakıldığında ise sistemin bit hata oranları anlamında iki farklı durumu birbirinden yine büyük oranda ayırmasıyla birlikte artık örneklerin ayrılma sınırı yakınlarına toplandığı ve birbirine daha yüksek oranda girişim yaptığı görülebilir.



Şekil 4.5 : Bütün saldırılar için BER

Şekil 4.5'teki genel dağılıma bakıldığında da sistemin bozulmuş videolar ve farklı videolar için ürettiği parmak izlerinin bit hata oranları kullanılarak ayrılabilceği görülebilir.

Bit hata oranları yukarıdaki şekilde gösterildikten sonra sistemin eşleme performansına göz atılabilir. Çizelge 4.3'te sistemin global ve geometrik saldırılar karşısındaki performansı bit hata oranları, doğruluk, hatırlama ve NDCR cinsinden verilmiştir.

Çizelge 4.3 : İçerik tabanlı video parmak izi sisteminin başarımı

		BER(Eşlik) (%)	BER (Fark) (%)	Doğruluk (%)	Hatırlama (%)	NDCR
Global	Ortalama	6,00	13,54	74,54	79,75	0,2025
	St. Sapma	2,48	2,48			
Geometrik	Ortalama	8,92	13,63	45,50	51,55	0,4845
	St. Sapma	2,12	1,72			

Çizelge 4.3 incelendiğinde, %80'e yakın doğruluk ve hatırlama oranları ile çıkarılan parmak izlerinin global saldırılara karşı dayanımının yüksek olduğu söylenebilir. Ayrıca 0-1 arasında değişen ve 0,2'nin civarında olunması hedeflenen NDCR değerine bakıldığında da parmak izi performanslarının global saldırılara karşı hedeflenen başarıma yaklaştığı görülebilir. Ancak yine aynı çizelgeye bakında, önerilen parmak izlerinin geometrik dönüşümlere karşı performansının düşük olduğu da görülebilir. Bunun önüne geçmek için parmak izi sisteminin bir önceki yöntemde olduğu gibi DD-NOMA veya benzeri bir geometrik dönüşümlere dayanımı artırmayı hedefleyen başka bir yöntemle desteklenmesi gerektiği söylenebilir.

5. SONUÇ VE ÖNERİLER

Bu çalışmada yayınlanan/dağıtılan video içeriklerinin izlenmesi (*video broadcast monitoring*) ve video kopyalarının belirlenmesi (*video copy detection*) gibi uygulama alanlarına sahip olan video parmak izi çıkarma problemi irdelenmiştir. Literatürde bulunan yöntemlerin aksine bu çalışmada video parmak izi çıkarma problemine negatif olmayan matris ayrıştırma (NOMA) yöntemi temel alınarak bir çözüm önerilmiştir. NOMA yönteminin bu problem için seçilme nedeni bu yöntemin verinin yerel özniteliklerini vurgulayabilmesi, veri özniteliklerini birbirinden bağımsızla yakın ayırabilmesi ve çok büyük boyutlu olan ham video verisinin çok küçük boyutlar ile temsil edilebilmesini sağlamasından gelmektedir.

NOMA yöntemi kullanılarak elde edilen parmak izlerinin TRECVID veritabanında denenmesi göstermiştir ki NOMA vektörleri hem anahtar-çerçeve tabanlı yöntem için hem de video tabanlı yöntem için global saldırılara karşı veriyi oldukça iyi temsil etmektedir. Önerilen yöntemlerin her ikisi de TRECVID başarı ölçütleri içerisinde kalan global saldırılara karşı oldukça gülbüz olmakla birlikte video verisi üzerindeki hafif miktarlardaki geometrik saldırılara da karşı koyabilmektedir. Ancak önerilen her iki parmak izi yönteminin de ciddi geometrik saldırılara karşı bir gülbüzlük sağladığını söylemek mümkün değildir. Bu yüzden ilk yöntem literatürde daha önceden var olan DD-NOMA yöntemi ile desteklenmiş ve parametreleri önceden bilinen geometrik dönüşümlere dayanıklı hale gelmesi sağlanmıştır. Bu yöntem ciddi geometrik ataklara karşı dayanıklılık sağlamakla birlikte, pratikte bir sistemde karşılaşılabilecek bütün saldırı türlerini önceden bilmenin mümkün olmaması bu yöntemin kullanılabilirliğini azaltır. Ayrıca DD-NOMA yöntemi sonucunda parmak izi elde edilmez ve eşleme basit bir parmak izi karşılaştırma adımı yerine döngüsel bir geri çatma işlemi sonucunda gerçekleşir. Bu da böyle bir sistemin gerçek zamanlı çalışmasını güçleştirir. İşte bu nedenle bu çalışmada anahtar-çerçeve tabanlı bir arama-getirme sisteminde doğrudan DD-NOMA kullanılması yerine, öncelikle Bölüm 3.1.1’de önerilen parmak izlerinin kullanılması, DD-NOMA’ya ancak gerek duyduğunda başvurulması önerilir. Böylelikle gerçek

zamanlı uygulamalarda gerek kalmadıkça geri çatma işleminin yükü altına girmekten kurtulunur. Ancak önerilen ikinci sistemde parmak izleri desteklemek için benzer bir sistem kurulmadığı için salt parmak izlerinin yetersiz olduğu bu durumda performansı artırmak için ileri bir aşamada DD-NOMA yönteminin anahtar-çerçeveler değil, doğrudan video verisi üzerinde işlem yapacak bir hale getirilmesi düşünülebilir.

Bunlara ek olarak Bölüm 3.2’de önerilen parmak izleri birden fazla ardışıl video çerçevesini aynı anda kullanması ile Bölüm 3.1’de anlatılan parmak izlerine göre video kliplerinin içeriğini daha iyi temsil edebilir. Çünkü bir anahtar-çerçeveden elde edilen bir parmak izi aslında önu ve ardındaki çerçevelerdeki verilere ilişkin bilgi içermemektedir. Bu durum özellikle hareketli sahnelerde anahtar-çerçeveli yöntemin sıkıntı çekmesine neden olur. Böyle bir sistemde, parmak izleri sadece anahtar çerçevelerden çıkarılmış olduğu için, gelen bir sorgu videosunun ilgili referans video klibi ile eşzamanlı olmaması gibi durumlarda (örneğin sorgu videosu referans video içinden çekilip alınmış ufak bir parça olduğunda sıkça karşılaşılabılır) parmak izleri aynı anahtar-çerçeveler için çıkmış olmayacak bu da eşlemede sıkıntı yaşanmasına neden olacaktır. Halbuki Bölüm 3.2’de önerildiği gibi elde edilen bir video parmak izi önünde ve ardındaki çerçevelere ilişkin bilgiyi de içerdiğinden ilk yöntemle göre bu eşlemeyi daha rahat yapacaktır. Buna ek olarak video parmak izleri bu çalışmada test edildiği gibi birbirleriyle örtüşmeyen video parçaları yerine birbirleriyle örtüşen parçalardan da elde edilebilir. Bu özellikle az önce bahsedilen zaman kaymalarında büyük bir avantaj oluşturabilir.

Ayrıca bir başka dikkat edilmesi gereken nokta, bu çalışmada video parmak izleri elde edilirken video çerçevelerinin alt bloklara ayrılmasının nedeni özellikle ikili parmak izlerinin yeganeliğinin artırılmasını sağlamaktır. Çerçeve verisinin tümünden kullanılması, özellikle boyut düşürmek için küçültüldükleri de düşünüldüğünde çerçevelerin içerisindeki detayların kaybolmasına dolayısıyla da birbirine benzer çerçeve ve çerçeve grupları için çok yakın parmak izlerinin elde edilmesine yol açmaktadır. Çerçevelerin alt bloklara ayrılması ve alt bloklar için çıkarılan parmak izlerinin de parmak izi vektörlerine eklenmesi ile birbirlerine benzer çerçeveler arasındaki detay farkların ortadan kaybolmaması ve böylelikle birbirinden daha uzak parmak izleri elde edilip yeganeliğin artması sağlanır.

Sonuç olarak bu çalışmada, negatif olmayan matris ayrıştırma (NOMA) tabanlı her iki parmak izi yönteminin de özellikle global bozulmalara karşı kullanılabilir olduğu gözlenmiştir. Ayrıca yine NOMA tabanlı olan DD-NOMA yönteminin bu sistemleri geometrik dönüşümlere karşı destekleyebileceği de görülmüştür. İlerideki çalışmalarda DD-NOMA yönteminin işlem yükünü hafifletmenin ve video verisi üzerinde doğrudan çalışır hale getirmenin ya da çıkardığı özniteliklerin parmak izlerinin içerisine katmanın yolu bulunduğu takdirde önerilen sistemlerin performansının daha da artması beklenebilir.

KAYNAKLAR

- [1] **Gursoy O., Günsel B., Sengor N.**, “Transform Invariant Video Fingerprinting by NMF,” Proc 13th Int. Conf. on CAIP, pp.452-459, 2009.
- [2] **TRECVID**, TREC Video Retrieval Evaluation,
<http://www.nlpir.nist.gov/projects/trecvid/>
- [3] **Lee S. and Yoo C. D.**, “Robust Video Fingerprinting Based on 2D-OPCA of Affine Covariant Regions,” Proc. IEEE ICIP, pp. 2156-2159, 2008.
- [4] **Massoudi A., Lefebvre F., et.al**, “A Video Fingerprint Based on Visual Digest and Local Fingerprints,” Proc. IEEE ICIP, pp. 2297-2300, 2006.
- [5] **Radhakrishnan R., Jiang W. ve Bauer C.**, 2009. A review of video fingerprints invariant to geometric attacks, Proc. of SPIE
- [6] **Lee D., Seung H.**, “Algorithms for non- negative matrix factorization,” Adv. Neural Inform. Process. Systems, vol. 13, pp. 556–562, 2001.
- [7] **Monga V., Mihçak M. K.**, “ 2007. Robust and Secure Image Hashing via Non-Negative Matrix Factorizations,” IEEE Transactions on Information Forensics and Security, Vol.2(3-1), pp. 376-390, 2007.
- [8] **Bucak S. S., Günsel B.**, " Video Content Representation by Incremental Non-negative Matrix Factorization, " Proc.IEEE ICIP, vol 2. pp. 113-116, 2007.
- [9] **Bucak S.S.**, 2008. Artımsal negatif olmayan matris ayrıştırma ile gözetleme türü videolarda arkaplan modelleme, İstanbul Teknik Üniversitesi Yüksek Lisans Tezi, 1-15
- [10] **Jolliffe I.**, 2000. Principle Component Analysis, Springer, 2nd ed. edition, New York.
- [11] **Guillamet D., Vitria J.**, 2002. Nonnegative matrix factorization for face recognition, Fifth Catalanian Conference on Artificial Intelligence, Castellon, Spain, 2002, 336-344.
- [12] **Fogel P., Young S.S., Hawkins D.M. ve Ledirac, N.**, 2007. Inferential robust non-negative matrix factorization analysis of microarray data, Bioinformatics, 23, 44–49.
- [13] **Cichocki A., Zdunek R., Amari S.**, 2006. New algorithms for non-negative matrix factorization in applications to blind source separation, IEEE International Conference on Acoustics, Speech and Signal Processing, Toulouse, France, 2006, 621-624.
- [14] **Monga V., Mihcak M.K.**, 2006. Robust image hashing via non-negative matrix factorizations, IEEE International Conference on Acoustics, Speech and Signal Processing, Toulouse, France, 2006, 225-228.

- [15] **Lee D.D. ve Seung, H.S.**, 2000. Algorithms for non-negative matrix factorization, *Proceedings of Neural Information Systems*, 942-948.
- [16] **Hoyer P.O.**, 2004. Non-negative matrix factorization with sparseness constraints, *Journal of Machine Learning Research*, 1457-1469.
- [17] **Pascual-Montano A., Carazo J.M., Kochi K., Lehmann D. ve Pascual-Marqui R.D.**, 2006. Nonsmooth nonnegative matrix factorization, *IEEE Transactions and Pattern Analysis and Machine Intelligence*, 42, 403-415.
- [18] **P.O. Hoyer**, 2002. Nonnegative Sparse Coding, *IEEE Workshop on Neural Networks for Signal Processing*, Martigny, Switzerland, 2002, 557-565.
- [19] **Li S.Z, Hou X., Zhang H. ve Cheng Q.**, 2001. Learning spatially localized parts-based representations, *IEEE Conference on Computer Vision and Pattern Recognition*, Hawaii, USA, 2001, 207-212.
- [20] **Lin C.-J.**, 2007. Projected gradient methods for non-negative matrix factorization, *Neural Computation*, 19, 2756-2779.
- [21] **Kim S. ve Dhillon I.S.**, 2007. Fast Newton-type methods for the least squares nonnegative matrix approximation problems, *Sixth SIAM Conference on Data Mining*, Minnesota, USA, 2007, 343-354.
- [22] **Smaragdis P., Brown J.C.**, 2003. Nonnegative matrix factorization for polyphonic music transcription, *IEEE Workshop Applications of Signal Processing to Audio and Acoustics*, New York, USA, 2003, 177-180.
- [23] **Li T. ve Ding C.**, 2006. The relationships among various nonnegative matrix factorization methods for clustering, *Sixth IEEE International Conference on Data Mining*, Hong Kong, 2006, 362-371.
- [24] **Shen J., Israel G.W.**, 1989. A receptor model using a specific non-negative transformation technique for ambient aerosol, *Atmospheric Environment*, 23(10), 2289-2298.
- [25] **Paatero P. ve Tapper U.**, 1994. Positive matrix factorization-a nonnegative factor model with optimal utilization of error estimates of data values, *Environmetrics*, 5, 111-126.
- [26] **Lee D.D. ve Seung H.S.**, 1999. Learning the parts of objects by nonnegative matrix factorization, *Nature*, 401, 788-791.
- [27] **Bucak S.S. ve Günsel B.**, 2007. Video content representation by incremental non-negative matrix factorization, *IEEE International Conference on Image Processing*, San Antonio, TX, USA, 2007, 113-116.
- [28] **Bucak S.S. ve Günsel B.**, 2009. Incremental subspace learning via non-negative matrix factorization, *Pattern Recognition* vol 42, no.5, pp. 788-798
- [29] **Eggert J., Wersing H. ve Körner E.**, 2004. Transformation-invariant representation and NMF, *Proc. of IEEE Int. Joint Conf. on Neural Networks*

- [30] **Fridrich J. ve Goljan M.**, 2000. Robust hash functions for digital watermarking, Proc. of ITCC
- [31] **FFmpeg**, Software Development Library, <http://www.ffmpeg.org/index.html>
- [32] **Bradski G.**, “The OpenCV Library”, Dr.Dobb’s Journal of Software Tools, 2000
- [33] **Cirakman O., Günsel B., Sengor N.N. ve Gursoy O.**, 2010. Key-Frame Based Video Fingerprinting by NMF, IEEE International Conference on Image Processing (ICIP)

ÖZGEÇMİŞ

Ad Soyad: Özgün ÇIRAKMAN
Doğum Yeri ve Tarihi: İstanbul / 19.06.1985
Lisans Üniversitesi: İstanbul Teknik Üniversitesi

Yayın Listesi:

- Ozgun Cirakman, Bilge Günsel, N.Serap Sengor ve Ozan Gursoy, “Key-Frame Based Video Fingerprinting by NMF”, *IEEE International Conference on Image Processing (ICIP)*, 2010