

# ÇOKLU ORTAM ONTOLOJİLERİNİ KULLANAN ANLAMSAL VIDEO ANALİZİ ÜZERİNE BİR İNCELEME

**Yasemin YÜKSEK ve Bahadır KARASULU**

Bilgisayar Mühendisliği Bölümü, Ege Üniversitesi, Üniversite Cad., 35100, Bornova, İzmir.  
[yasemin.yuksekk@ege.edu.tr](mailto:yasemin.yuksekk@ege.edu.tr), [bahadir.karasulu@ege.edu.tr](mailto:bahadir.karasulu@ege.edu.tr)

(Geliş/Received: 08.12.2009; Kabul/Accepted: 24.03.2010)

## ÖZET

Çoklu ortam belgeleri; görüntü, metin, ses, video gibi verilerden oluşur. Bunların analiz edilmesi, bunlara ait verilerden bilgi elde edilmesi literatürde araştırmacılar tarafından oldukça yoğun bir biçimde uğraşılan konulardandır. Video akışları, birkaç çoklu ortam bileşenini içeren oldukça karmaşık çoklu ortam verileridir. Bu verilerin analizinde sözdizimsel yapının yanısıra anlamsal yapılar da kullanılmaktadır. Video endeksleme, özetleme konularında ontolojilerin kullanımı gün geçtikçe artmaktadır. Video olayları basit, bileşik ve karmaşık olaylar olarak sınıflandırılabilirler. Bu tespitlerden yola çıkarak, inceleme çalışmamızda Anlamsal Web için literatürde varolan ontolojiler, çoklu ortam ontolojileri ve bunların literatürdeki örnekleri detaylı olarak incelenmiştir. Video olay analizi konusu ve anlamsal video olay analizi için ontolojilerin kullanımı konuları incelenerek, literatürdeki örnek çalışmalar üzerinden değerlendirme yapılmıştır.

**Anahtar Kelimeler:** Ontoloji, Çoklu ortam ontolojileri, Video anlamsal analizi, Video içerik-tabanlı bilgi elde etme.

## A REVIEW ON SEMANTIC VIDEO ANALYSIS USING MULTIMEDIA ONTOLOGIES

### ABSTRACT

Multimedia documents involves some data such as images, text, audio and video. In literature, the information retrieval from multimedia documents via multimedia analysis is an intensive research issue for researchers. Video streams are very complex multimedia data which covers some other multimedia component. In analysis of these data, it can be used both syntactic and semantic structures. Usage of ontologies on video indexing and abstraction issues are increase more and more. The video events can be classified such as simple, composite and complex events. As a results of those, we reviewed in our review study in details the existing ontologies and multimedia ontologies for the Sematic Web and their examples in literature. Also, we reviewed video event analysis and usage of ontologies for semantic video event analysis, and we made a conclusion over example works in literature.

**Keywords:** Ontology, Multimedia Ontologies, Video semantic analysis, Video content-based retrieval.

### 1. GİRİŞ (INTRODUCTION)

Günümüzde çoklu ortam uygulamaları, hızla gelişim gösteren araştırma alanıdır. Gelişim en çok etkileşimli TV, sayısal video yayıncılığı, istek anında video, sayısal kütüphaneler ve uzaktan eğitim alanlarında olmaktadır. Birçok yeni teknik çoklu ortam verilerinin organizasyonu ve soyutlanmasında önemli rol oynamaktadır. Bu konuda karşılaşılan problemlerden birisi, çoklu ortam verisinin nasıl en etkin yoldan

özetlenebileceğidir. Özellikle video veritabanları ve uygulamalarında karşılaşılan problemlerin çözümü için sayısal video akışlarının analiz edilebilmesi gerekir. Video veritabanlarının endekslenmesindeki ilk aşama olan video akışının analiz edilmesi işlemi iki önemli adımı içermektedir. Birinci adım, videonun anlamlı ve yönetilebilir birimlere bölünmesidir. İkinci adım ise, bu birimler içerisinden anahtar çerçeve çıkarımı olarak bilinen adımdır [1].

Çoklu ortam verilerinin anlamsal analiz problemine çözüm olarak birçok yaklaşım önerilmiştir. Bu yaklaşımlar, düşük-seviyede sayısal görsel-işitsel (audio-visual) verisi ile yüksek-seviyede sadece insanın anlayabileceği kavramlar ve elemanlar arasındaki anlamsal boşluk için köprü görevi görecek içerik-tabanlı anlamsal sınıflandırma yaklaşımlarıdır. Anlamsal Web [2-5], makinelerin ve insanların birlikte çalışmasına imkân veren bugünkü web'in uzantısıdır. Video içeriğinin anlamsal etiketlenmesi, video içerik alanını modellemede kullanılan ontolojiler ile gerçekleştirilmektedir. Anlamsal Web'in temel bileşeni ontolojiler [8-10], çoklu ortam veri yapısını ve çoklu ortam verileri arasındaki ilişkileri anlamsal tanımlamalar ile ifade etmektedirler [6, 7]. Çoklu ortam modellemenin karmaşıklığı nedeniyle Çoklu ortam bilgilerinin modellenmesi için, Çoklu ortam Web Ontoloji Dili (Multimedia Web Ontology Language - M-OWL) kullanılmaktadır.

Çalışmanın ikinci bölümünde çoklu ortam içerik analizi ve detayları ile video analizi anlatılmaktadır. Üçüncü bölüm içerisinde web ontolojileri ve çoklu ortam ontolojileri, çoklu ortam anlamsal dilleri ve videolarda olay, anlamsal video olay tanımlama ve dilleri anlatılmaktadır. Dördüncü bölümde ise literatürdeki önceki çalışmaların sonuçları hakkında değerlendirme ve tartışma yer almaktadır.

## 2. ÇOKLU ORTAM İÇERİK ANALİZİ (MULTIMEDIA CONTENT ANALYSIS)

Çoklu Ortam İçerik Analizi (ÇOİA), geçtiğimiz yıllarda önemini oldukça arttırmış bir konudur. Bu nedenle, görsel-işitsel veri akışlarından anlamsal içerik bilgisi çıkarımı için birçok yöntemler geliştirilmiştir. ÇOİA'nın karşılaştığı sorunların başında, düzgün biçimde veri akışlarının zamansal kesimlenmesinin yapılabilmesi gelmektedir. Bu sayede video akışları gibi oldukça karmaşık yapılar, video çekimleri (shot), sahneler (scene), olaylar (event) gibi bölümlere ayrılabilen ve anlamsal endeksleri kullanarak oluşturulan zamansal bölümlerin etiketlenmesi ve sınıflandırılmasına da olanak sağlanmaktadır. Genellikle, bir ÇOİA algoritması önceden etiketlenmiş bir test veri setinin *ground truth* (yani 'altın standart' olarak bilinen ve bir uzman insan tarafından üzerinde işaretleme yapılmış referans veri) olarak ele alınması yoluyla değerlendirilir [11].

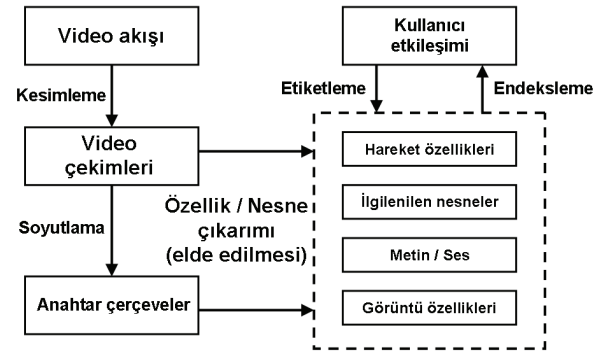
Videoda olaylar genellikle basit, bileşik ve karmaşık olaylar olarak ele alınmaktadır. Basit olaylar, videodaki tek bir nesnenin yaptığı hareketlerle ifade edilirken, bileşik olaylar birden çok nesnenin aralarındaki etkileşimle ifade edilmektedirler [12].

### 2.1. Anlamsal Video Analizi (Semantic Video Analysis)

Videolardan içerik tabanlı bilgi elde edilmesinin en etkin yolu, video verisinin düzgün bir biçimde ifade edilmesiyle olabilmektedir. Bir bütün video akışı bazı

yapısal birimlerden oluşur. Bunlar sırasıyla; sahne, video çekimleri ve çerçevelerdir. Video çekimleri, bir videoda genel bir sahneyi paylaşan ve tek bir kamera tarafından çekilen ardışık çerçevelerden oluşur. Video çekimleri arası geçişler çeşitlidir, kolay-tespit edilebilen kesin kesitler veya zor çözümlenebilen daralmalar, kararmalar gibi geçişleri içerebilmektedir. Bu tarz geçişler iki ana gruba ayrılırlar: Kırıklı geçiş ve Kademeli geçiş. Kırıklı geçişler aniden oluşur ve bir çerçeve önceki video çekimine ait iken sonraki çerçeve başka bir video çekimine ait olur. Kademeli geçişlerde ise iki video çekimi sanki harmanlanmış gibi içiçe bir geçiş sağlamaktadır [13].

Bilgi elde etme teknikleri dört ana grupta toplanabilir: Video çekimi sınırı tespiti, anahtar çerçeve seçimi, anahtar çerçevelerden düşük-seviyeli özellik çıkarımı ve bilgi elde edimi. Anahtar çerçeve seçimi; her bir video çekimi içerisinde çeşitli özellikleri içeren belirleyici bir çerçevenin seçilmesi sayesinde gerçekleştirilebilen bilgi özetlemesidir. Anahtar çerçevelerden düşük-seviyeli özellik çıkarımı ise; renk, doku, biçim ve nesnelerin hareketleri gibi temel özelliklerin çıkarımı, anahtar çerçevelerin ve daha sonrasında da video çekimlerinin belirlenmesi için çıkarım işlemidir. Bilgi elde etme işlemi, son kullanıcı tarafından oluşturulan girdi biçimindeki bir sorgunun, belirli bir veritabanı üzerinde yapılacak arama yoluyla belirli bilginin simetrik eşlenmesi sonucu döndürülecek cevapla gerçekleştirilmektedir [14]. Şekil 1'de, kavramsal bir içerik-tabanlı video endeksleme ve arama sistemi gösterilmektedir [15].



Şekil 1. Video içeriklerini ayrıştırma, sunma ve organize etme aşamalarını içeren kavramsal sistem. (The conceptual system which involves video content parsing, content presentation and organization stages)

Günümüzde önemli sayıda video, çeşitli kaynaklardan oluşturulmakta ve sayısal kütüphanelere eklenmektedir. Bu içeriğin yönetilmesi için etkin bir depolama, endeksleme ve bilgi elde edilmesi önemlidir. Güncel veri modelleri, video veritabanları üzerinde uzamsal-zamansal (spatio-temporal) sorgulamayı desteklemektedirler [16-19]. Kural-tabanlı video veritabanı sistemlerine örnek olarak; 2002 yılında Dönderler vd. [16] çalışmalarında hazırladıkları sistem, uzamsal-zamansal ve anlamsal sorgularını desteklemektedir. Ayrıca 2004 yılında,

Köprülü vd. [17] çalışmalarında sorgular, bulanık uzamsal-zamansal ilişkiler şeklinde biçimlendirilmiştir. Bu ilişkiler, hareketli nesnelerin rotaları ve video nesneleri arasındaki ilişkilerdir. Bunun yanı sıra, İçerik-tabanlı Video Sorgulama Dili (Content-based Video Query Language - CVQL) isimli bir sorgulama dili sunulmuştur. CVQL sayesinde uzamsal-zamansal ilişkileri içeren tahmin ile sorgulama yapılabilmektedir [19].

Jung ve Park [20] 2008 yılındaki çalışmalarında, bir sahne birimini temel alan video verisinden bilgi elde etmek için video ontolojisi önermişlerdir. Önerilen sistem, videodan bilgi elde etmenin temel birimi olarak bir anlamsal sahne yaratmaktadır.

### 2.1.1. İçerik-tabanlı Video Endeksleme ve Bilgi Elde Etme (Content-based Video Indexing and Information Retrieval)

İçerik-tabanlı video endeksleme ve bilgi elde etmek için dört ana süreç bulunmaktadır: Video içerik analizi, video yapısının ayrıştırılması, video özetleme veya soyutlama ve video endeksleme. Bunlar sırasıyla aşağıdaki maddeler halinde özetlenmiştir.

Video içerik analizi: sürecindeki ana problem; görsel özelliklerin (renk, doku, biçim, yapı, hareket vs.) anlamsal kavramlara (iç veya dış mekân, insan veya araba yarışı sahneleri) kolaylıkla bağdaştırılmamasıdır. Bunun yanı sıra görsel içerik bir videodaki bilgi kaynağıdır. Bu tarz yaklaşımlara örnek olarak; Informedia sistemi [21] veya AT&T'e ait Pictorial Transcripts sistemi [22, 23] verilebilir.

Video yapısının ayrıştırılması: videoyu tek tek sahnelere bölümlenmede kullanılır. Bir sahne aynı tematik içeriğe sahip çeşitli ardışık video çekimlerinden oluşmaktadır [23].

Video özetleme: video yapısı ile ilgili görsel bilgi gösteriminin oluşturulması için bir süreçtir. Bu soyutlama süreci, metin belgelerindeki özetlerin veya anahtar kelimelerin çıkarılmasına benzemektedir [23].

Video endeksleme: yapısal ve içerik nitelikleri; içerik analizinde, video ayrıştırmada ve özetleme süreçlerinde bulunmaktadır. Bu tarz nitelikler genellikle üstveri (metadata) olarak adlandırılır. Bu nitelikler temel alınarak; video indisleri ve/veya içerik tabloları oluşturulabilir. Diğer bilgi sistemlerinde olduğu gibi, bu sistemlerde de video indisleri ve içerik üstverilerin kullanımı ile büyük video veritabanlarının sorgulanması, aranması ve incelenmesi işlemlerinde şemalar ve araçlar kullanılmaktadır [23].

Bunların haricinde; sorgulamalar, görsel ve kavram sorgulama olarak ikiye ayrılır. Görsel sorgulama ile kullanıcı, verilen örnek ile aynı görünüme sahip video çekimlerini bulmak için sorgulama yapmaktadır. Kavram sorgulamada ise kullanıcı, özel bir nesneyi

veya olayı içeren video çekimlerini bulmak için sorgulama yapmaktadır [15].

Petkovic ve Jonker 2001, 2002, 2003 [24-26] yıllarındaki çalışmalarında, içerik-tabanlı video bilgi elde etme sistemlerindeki anlamsal boşluğu doldurmaya uğraşmışlardır. Bunu yapmak için, uzamsal-zamansal somutlaştırma tabanlı Gizli Markov Modelleri (Hidden Markov Models - HMM), Dinamik Bayes'çi ağlar gibi tekniklerin kullanımını önermişlerdir. Ayrıca Petkovic ve Jonker [26] çalışmalarında, İçerik-tabanlı Video Bilgi Elde Etme Sistemi'ni (Content-based Video Retrieval System - COBRA) önererek, video verisini dört katmanlı olarak incelemişlerdir. Bu katmanlar sırasıyla, ham (raw) video verisi, özellik, nesne ve olay katmanlarıdır. Bu sistem, ham video verisinden özellik çıkarımı ve renkli histogram-tabanlı video çekimi tespitini desteklemektedir [19].

### 2.1.2. Anlamsal Video İçerik Sınıflandırma (Semantic Video Content Classification)

Literatürdeki çalışmalarda genellikle video içerikleri kavramsal olarak iki parçaya ayrılırlar. Bunlardan ilki anlamsal içeriktir. Anlamsal içerik video cinsi, olaylar ve nesneler olarak alt gruplara ayrıştırılabilir. Bu grup üyeleri tüm ortamlarda da bulunabilirler. Örneğin bir kitabın korku veya bilim-kurgu tarzında olması onun "cinsini", içinde anlattığı hikâyelerin ilerleyiş aşamaları "olayları" ve "nesneleri" ifade etmektedir. İkinci parça ise düzenleme olarak ele alınabilir. Buna göre literatürdeki çalışmalarda, sayısal ortam videolarının değişmez bazı özellikleri olduğu ve bunlara ait düzenleme etkilerinin de olduğu belirtilmektedir [27].

Video cinsi (genre) sınıflandırma: bir videonun tarzını yansıtan olguyu videonun cinsi olarak tarif etmek ilk aşamada onu anlamsal olarak sınıflandırma anlamına da gelmektedir. Video cinsi tanımlama, video içeriğinin izlenmesi ile oluşturulmaktadır. Genellikle nesnel olmayan görüş ve anlamsal farklılıklar sürece dâhil olur. Fischer vd. [28], 1995 yılındaki çalışmalarında, video cinsi sınıflandırmayı denemişlerdir. Sınıflandırmaları arasında; haber, reklam, çizgi-film, araba yarışı gibi video cinsleri bulunmaktadır. Liu vd. [29, 30] çalışmalarında da televizyondan elde edilmiş videolar için video cinsi sınıflandırma yaklaşımını denemişlerdir. İlk çalışmalarında bir sinir ağı sınıflandırıcı kullanmışlardır ve aynı test düzeneğine HMM uyguladıklarında %11,9 iyileşme gözlemlediklerini belirtmişlerdir. İkinci çalışmalarında %93,4 doğrulukta video cinslerini üç ayrı sınıfa ayırtırmayı başarmışlardır, bunlar sırasıyla spor, raporlar, reklamlar olarak verilmiştir [27].

Videolarda olay (event) sınıflandırma: video içerisindeki anlamsal içeriklerden biri olan olayların sınıflandırılması video cinsi sınıflandırılmasındakine

benzer bir yol izlenmesi ve olayların bazılarının diğer olaylardan daha önemli olması, olayların sınıflandırılması gerekliliğini ortaya koymaktadır. Chang ve Lee [31] çalışmalarında, bir voleybol karşılaşmasına ait videodaki olayları sınıflandırmışlardır. Colombo vd. [32] çalışmalarında, renk ve hareket özelliklerini, eylem, mutluluk, gevşeme gibi anlamsal özellikleri oluşturmada kullanmışlardır. Olaylar çoğu çalışmada karşılıklı dışlayan (mutually exclusive) biçimde tanımlanmışlar. Burada karşılıklı dışlama anlam olarak, “iki farklı olay verildiğinde bunlardan birinin gerçekleşmesinin, diğerinin gerçekleşmesini olanaksız kılmasıdır” şeklinde ifade edilir [27].

**Video nesne (object) sınıflandırma:** olaylar çeşitli sayıda ve biçimdeki nesnelere içerebilmektedir. Bunlar kavramsal olarak, video içeriğinin anlamsallığına etki eden sınıflandırmanın en düşük-seviyesi olan birimlerdir. Literatürde en çok tespit edilmeye ve sınıflandırılmaya uğraşılan nesne insan yüzüdür [33, 34]. Nesne tespiti çoğunlukla iyi bir biçimde yapılandırılmış özellik çıkarımını gerektirir. Yapı içerisinde oldukça iyi bir biçimde tanımlanmış durumdadırlar. Kural-tabanlı algoritmalar sayesinde özellik çıkarımı yapılabilen ve bu yolla tespit edilen tanımlı nesnelere arasındaki ilişkiler ve etkileşimler anlamsal olarak olayların tespiti ve sınıflandırılmasında kullanılabilir [27].

### 2.1.3. Anlamsal Olay Analizi (Semantic Event Analysis)

Literatürde, videolardaki olaylara ait bir terminoloji üzerinde ortak bir karara henüz varılabilmemiş değildir. Her yeni çalışma kendisine özgü farklı terimleri de beraberinde getirmektedir. Örneğin, Hu vd. [35] çalışmalarında, “nesne davranışları” na dikkati çekmişler ve bunları zamanla-değişen veri olarak tanımlamışlardır. Aggarwal ve Cai [36] çalışmalarında, “aktivite/davranış” kavramını bir görüntü dizisinden elde edilmiş örüntü olarak tarif etmişlerdir. Cohn vd. [37] çalışmalarında, “dinamik davranış” zamansal mantık tarafından tanımlanmış ilişkileri bulunan uzamsal durumların bir kümesi olarak tanımlamışlardır. Hongeng ve Nevatia [12] çalışmalarında, “basit olayları” doğrusal bir zaman dizisinde oluşan eylemler olarak tanımlamışlardır.

Olay modelleme, video soyutlamanın temel problemlerinden birisidir. Bu yolla, olayların belirli bir alanda tanımlanması işlemi, verilen bir soyutlama şeması ile yapılabilmektedir. Belirli bir olay modeli, belirli alandaki belirgin olayların gösterimindeki kapasitesi ve video girdi dizisindeki bu olayların oluşlarını tanımadaki kapasitesine bakılarak seçilir. Lavee vd. [38], 2009 yılındaki çalışmalarında da belirtildiği gibi, olay modelleri; deterministik ve olasılıksal olay modelleri olarak iki ana kategoriye ayrılabilirler. Buna göre deterministik modellerde

olayların sırası ve süresi deterministik olurken, olasılıksal modellerde olay sırası ve süresi olasılık durumlarının içerildiği küme şeklinde belirtilmektedir.

**Tekil-altsüreç (Single-thread) olayları:** basit bir olay, bir gözlem sisteminin gözlemediği alan içindeki herhangi bir nesnenin belirli bir zaman aralığındaki hareketi olarak alınabilir, bu zaman zarfında bu nesneye ait belirli özelliklerin ardışık çevreler boyunca gözlemlenmesi de beklenmektedir.

Bileşik olaylar, uzamsal-zamansal olarak ilişkili olan basit olaylar kümesi olarak ifade edilirler. Micheloni vd. [39], 2009 yılındaki çalışmalarında da belirtildiği gibi; karmaşık olaylar, önceden tanımlı bazı örüntüleri takip eden ve zamansal olarak sıralı olan basit olaylar dizisinin bulunması anlamına gelmektedir. Bunun başarılabilmesi için çizge eşleme oluşturmak gerekmektedir. Çizge eşleme, klasik eniyileme yaklaşımıdır.

**Çoklu-altsüreç (Multi-thread) olayları:** aralarında aynı mantıksal ve zaman ilişkisine sahip iki veya daha fazla tekil-altsüreç olayından sorumludur. Her bir bileşik altsüreç, farklı aktörler tarafından yerine getirilebilir. Bu durumlarda çoklu-altsüreç olayları, nesnelere arasındaki etkileşimleri modellemek için göz önüne alınabilirler. Oliver vd. [40], 2000 yılındaki çalışmalarında, çiftlenmiş HMM’leri nesnelere arasındaki etkileşimleri modellemekte kullanmışlardır.

### 2.1.4. Başarım Ölçümü ve Değerlendirme (Performance Measuring and Evaluation)

Geçtiğimiz yıllardaki birçok başarım ölçüm ve değerlendirme projelerinden en önemlilerinden birisi de metinden bilgi elde edilmesi işleminin değerlendirilmesinin hedeflenerek her yıl düzenlenen bir konferans serisine ek kısım olarak 2003 yılında “Videodan bilgi elde etmenin değerlendirilmesi (Text Retrieval Conference–Video Retrieval Evaluation Track - TRECVID)” isimli yeni bir değerlendirme projesi haline gelmiş şeklidir. TRECVID projesinde, verilen test kümesi için en iyi video bilgi elde etme sistemini geliştirmeye çalışmaktadırlar [41]. Burada karşılaşılan en büyük zorluk ise, belirli bir tip için test yığınlarını belirli tipte kullanıcı ile bir araya getirmek ve videolardan alınmış video çekimleri için verilecek ilgili kararların kümesini de buna dâhil etmekte yaşanan sıkıntılardır. TRECVID karşılaştırmalı sinema çalışmaları Amerikan Ulusal Standartlar ve Teknoloji Enstitüsü (National Institute of Standards and Technology - NIST) tarafından yürütülmektedir. Bu çalışmalar, anlamsal kavram tespitindeki en güncel tekniklerin başarım değerlendirmeleri için kullanılmaktadır [42]. Bu çeşit tekniklerin detaylı bir incelemesi ve kavram tespiti için etkinlikleri Naphade vd. [43] çalışmalarında verilmektedir. TRECVID-2005 proje çalışmasında, otuzdokuz adet kavram

başarım değerlendirmesi için sağlanmaktadır. Bunlar arasında: İnsan, Yürüme/Koşma, Yangın, Harita, Amerikan bayrağı, Bina, Dış Mekân, Deniz Manzarası/Rıhtım, Dağ, Spor, Araba vs. olarak sayılabilir.

TRECVID-2006 [44] veri seti 61901 adet eğitim video çekimi ve 79484 adet test video çekimi içermektedir. TRECVID-2007 [44] veri seti ise bilim haberleri, haber bültenleri, belgeseller gibi çoklu ortam belgelerini içermektedir. Yüz saati aşkın video verisi, otuzaltı farklı tanımlı etikete sahip anlamsal kavramlar video çekimlerine ayrıştırılmış ve etiketlenmişlerdir [45].

Başarım için en önemli konu, karşılaştırmalı sına testlerinin ötesinde başarım grafiklerinin tamamlanmasıdır [46]. Bir bilgi elde etme sisteminde, birçok çalışmada en kullanışlı başarım ölçümleri duyarlık (precision), geri çağırım (recall) ve ortalama duyarlık (average precision) ve bazı çalışmalarda da duyarlık ortalamalarının ortalamasıdır. Huijsmans ve Sebe [46] çalışmalarında duyarlık-geri çağırım grafiklerinin sınırlılıklarını anlatmışlardır.

Anahtar çerçeve tabanlı bilgi elde etme teknikleri, video bilgi elde etme sistemleri içerisinde en popüler olanlardır. Bu sistemler bir videoyu, video içeriğinden alınan çerçevelerin küçük bir kümesi olarak gösterirler. Pickering ve Rüger [47] çalışmalarında, iki adet öğrenme yöntemine (bunlar destekleme (boosting) ve k-ortalama yöntemleridir) karşılık bir vektör uzay modeli ile oluşturdukları değerlendirmeyi önermişlerdir. Silva vd. [48] çalışmalarında birkaç anahtar çerçeve çıkarımı algoritması gerçekleştirmişler ve kamera değişimleri tabanlı uyarlanabilir bir algoritmanın kullanılmasının yüksek kalitede sonuçlar verebileceğini bulmuşlardır.

Çoklu ortam bilgi elde etme değerlendirmesi literatürde sıkça incelenmektedir [41, 46, 49, 50]. Video gözetleme sistemleri için başarım değerlendirme konusunda oldukça fazla proje ve çalıştay bulunmaktadır. Kolaylıkla elde edilebilen etiketlenmiş video dizileri birçok başarım değerlendirme projesinde mevcuttur. Bunlar arasında CAVIAR [51], VACE [52], CREDS [53], CLEAR [54] projeleri ve PETS [55] çalıştay sayılabilir. Bu araştırma programları video dizilerini çeşitli genel zorluk seviyelerinde ilişkilendirilmiş *ground truth* ile bir arada sağlarlar. Sonuç olarak, değerlendirme süreci her bir video işleme algoritmasında aynı sonuçlara ulaşmamaktadır. ETISEO [56], algoritmalar ile videoya ait yapısal karakteristikler arasındaki bağımlılığı ortaya çıkartmaya yoğunlaşmış bir başarım değerlendirme projesidir.

Video etiketlemede ölçüm ve değerlendirme: genel bir video etiketleme başarım ölçümü ve değerlendirmesi

için duyarlık (precision), geri çağırım (recall) ve ortalama duyarlık ölçümleri şöyle tanımlanır:

$$\text{Duyarlık} = \frac{\{BV\} \cap \{EV\}}{\{EV\}} \quad (1)$$

$$\text{Geri çağırım} = \frac{\{BV\} \cap \{EV\}}{\{BV\}} \quad (2)$$

$$\text{Ortalama duyarlık} = \frac{\sum_{i=1}^m \text{Duyarlık}(r)\delta(r)}{\{BV\}} \quad (3)$$

Burada BV; benzer videolar, EV; elde edilmiş videolar olarak kullanılmıştır. Burada,  $r$  rankı'nın ilişkilğine bağlı ikili  $\delta$  fonksiyonu bulunur. Elde etme sıralamasına ait kalite, ortalama duyarlık değeri ile ayarlanabilmektedir. Bu ölçüt ile ilişkili olan videolar daha yüksek sıradaki konuma gelirler. Yani duyarlık arttıkça sorgu veya aramanın sonucu elde edilmiş video, istenilen videoya veya içeriğe çok benzer olan bir video olacaktır.

Video olaylarının tespitinde ölçüm ve değerlendirme: video olaylarının tespiti ve tanınması önemli bir anlamsal analizdir. Bununla ilgili başarım ölçüm ve değerlendirme; kullanılacak yöntemin geçerliliği ve doğruluğunu ortaya koymaktadır. Literatürdeki çalışmalarda, genel anlamı ile duyarlık ve geri çağırım değerleri başarım ölçümünde kullanılmaktadır. En sık kullanılan formülasyonlarda “Kaçırılan” terimi bir yanlış negatif, amaç olayının amaç dışı olay olarak yanlışlıkla sınıflandırılmış olmasıdır. “Yanlış tanımlanmış” terimi ise bir yanlış pozitif göstermektedir. Amaç dışı bir olayın amaç olay olarak tanımlanmış olmasıdır. Bunlara ait duyarlık ve geri çağırım formülleri aşağıdaki gibi verilmektedir.

$$\text{Geri çağırım} = \frac{\text{Tanımlanan}}{(\text{Tanımlanan}) + (\text{Kaçırılan})} \quad (4)$$

$$\text{Duyarlık} = \frac{\text{Tanımlanan}}{(\text{Tanımlanan}) + (\text{Yanlış tanımlanmış})} \quad (5)$$

Literatürdeki çeşitli çalışmalara bakıldığında, sorgulama başarım ölçümleri için Eşitlik 4 ve Eşitlik 5'e benzer yapıda belirli ölçütler olduğu görülmektedir. Fakat çalışmamızın konu kapsamı nedeniyle, “video etiketleme” ve “video olay tespiti” başarım ölçümlerine çalışmamızda yer verilmiştir.

SanMiguel vd. [57], 2009 yılındaki çalışmalarında, önerdikleri görsel analiz çatısının başarımını deneylerle ölçmüşlerdir. Deneylerini i-Lids [58] ve PETS2006 [55] veri setleri ile gerçekleştirmişlerdir. Önerdikleri olay altkümesinin tespitine ait sonuçlar için duyarlık ve geri çağırım değerlerini hesaplamışlardır. Başarım değerlerinin geleneksel yaklaşımlara göre daha düşük çıkmasını ontolojideki

olay tanımlamalarının daha çok olmasına bağlamışlardır. Çalışmalarında elde ettikleri değerler Tablo 1’de verilmektedir.

**Tablo 1.** SanMiguel vd. [57] çalışmasındaki olay tespiti için duyarlık ve geri çağırım sonuçları. (Precision and recall results for event detection in SanMiguel et al. [57] studies)

| Olay             | Elle etiketleme | Duyarlık | Geri çağırım |
|------------------|-----------------|----------|--------------|
| Nesne al         | 20              | %79      | %67          |
| Nesne yerleştir  | 33              | %77      | %60          |
| Bırakılmış nesne | 12              | %74      | %59          |
| Çalınmış nesne   | 8               | %75      | %62          |

Chen vd. [59], 2007 yılındaki çalışmalarında, geliştirdikleri algoritmalarının test edilmesi amacıyla toplam dokuz saat süren yirmiyedi adet futbol (soccer) videosunu kullanmışlardır. Çalışmalarında amaç olay tespiti ile ilgili başarımların değerleri olarak duyarlık ve geri çağırım değerlerini bir karşılaştırmada kullanılmasına olanak tanımaktadırlar. Çalışmalarında elde ettikleri değerler Tablo 2’de verilmektedir.

Bertini vd. [60], 2007 yılındaki çalışmalarında, futbol alanında ontoloji oluşturulması süreci bir eğitim kümesi kullanılarak yapılmıştır. Buna göre; oyun için 8 görsel kavram, taraftarlar için 3 görsel kavram, oyuncuların yakın plan çekimleri için 48 görsel kavram ve oyuncuların orta-seviyeli çekimleri için 20 görsel kavram verilmiştir. Bu ontoloji, deneylerde kullandıkları videoların otomatik olarak etiketlenmesi için kullanılmıştır. Bu deneylerde geliştirdikleri sistem ile olaylardan genel kavram tespiti ve tanımda elde ettikleri başarımların sonuçları Tablo 3’de verilmektedir.

Tablo 3’de görülebileceği gibi, başarımların ölçümleri

**Tablo 2.** Chen vd. [59] çalışmalarındaki algoritmaya ait amaç olay tespiti için duyarlık ve geri çağırım sonuçları (Precision and recall results for goal event detection of algorithm in Chen et al. [59] studies).

| Test No | Amaç Sayısı | Tanımlanan olay | Kaçırılan olay | Yanlış Tanımlanan olay | Duyarlık | Geri çağırım |
|---------|-------------|-----------------|----------------|------------------------|----------|--------------|
| Test 1  | 11          | 10              | 1              | 2                      | %90      | %83,3        |
| Test 2  | 11          | 11              | 0              | 2                      | %100     | %84,6        |
| Test 3  | 10          | 10              | 0              | 2                      | %100     | %83,3        |
| Test 4  | 12          | 11              | 1              | 2                      | %91,7    | %84,6        |
| Test 5  | 11          | 11              | 0              | 2                      | %100     | %84,6        |

**Tablo 3.** Bertini vd. [60] çalışmalarındaki yaklaşıma ait olaylardan genel kavram tespiti ve tanıma için başarımların sonuçları (Performance results for common concept detection and recognition from events of Bertini et al. [60] studies’s approach).

| Kavramlar         | Kaçırılan | Bilinmeyen | Yanlış | Duyarlık | Geri çağırım |
|-------------------|-----------|------------|--------|----------|--------------|
| Oyun              | %1        | %0         | %12    | %90      | %99          |
| Taraftar          | %10       | %0         | %20    | %82      | %90          |
| Yakın plan çekim  | %10       | %43        | %1     | %98      | %48          |
| Orta-seviye çekim | %16       | %18        | %16    | %79      | %65          |

olan duyarlık, geri çağırım ve ortalama duyarlık kullanımı ve bunların olay analizinde kullanımında da değer olarak alınan “Kaçırılan” ve “Yanlış tanımlanan” unsurları sıklıkla literatürde kullanılmaktadır.

### 3. ANLAMSAL WEB (SEMANTIC WEB)

Anlamsal Web, kullanıcı etmenleri ve diğer web tabanlı uygulamaların, iyi tanımlanmış bilgilerin ve servislerin kolay bir şekilde makineler tarafından anlaşılabilir olmasını sağlamaktadır. Tim Berners-Lee tarafından önerilen Anlamsal Web katmanlı mimari önerilmiştir. Bu mimari katmanları arasında aşağıdan yukarıya karmaşıklık seviyesi artmaktadır.

Günümüzde web ortamında düşük-seviyede görsel-işitsel veri ile yüksek-seviyede sadece insanın anlayabileceği kavramlar ve elemanlar arasında anlamsal bir boşluk oluşmaktadır. Bu anlamsal boşluğun, ontolojiler kullanılarak doldurulması amaçlanmaktadır. Böylece web ortamında bulunan çoklu ortam uygulamalarındaki bilgi kaynakları, tekrar kullanılabilir ve standart hale getirilecektir [8].

#### 3.1. Ontolojiler (Ontologies)

Ontoloji [9, 61, 62, 63], bir alanın açık ve biçimsel kavramsallaştırılmasıdır. Video içeriğinin anlamsal etiketlenmesi, video içerik alanını modelleyen alan ontolojileri kullanılarak yapılmaktadır. Çoklu ortam uygulamaları için tasarlanmış ontoloji, kavramsal ve ortam alanlarının birleştirilmesini sağlar.

Çoklu ortam verisinin anlamsal işlenmesinde, görsel-işitsel veriler ve soyut kavramsal veriler arasında bir köprü gerekmektedir. Bu nedenle kavramlarla ortam özelliklerinin açık belirtimi ve kavramların elde edilmesi için ontoloji tanımlama dilinin bir uzantısı olarak, M-OWL [8, 64] geliştirilmiştir. Alan ontolojisindeki kavramlar ve ilişkilerin ifade edilmesinde tanımlama dilleri olarak: Web Ontoloji

Dili (Web Ontology Language - OWL), Kaynak Betimleme Çatısı (Resource Description Framework - RDF), RDF Şeması (RDF Schema - RDFS), Hareketli Görüntü Uzmanlar Grubu - 7 standardı (Moving Picture Experts Group - 7 standard - MPEG-7) kullanılmaktadır [7].

### 3.1.1. Ontoloji Tanımlama Dilleri (Ontology Definition Languages)

Ontoloji tanımlama dilleri, alan modellerinin açık ve biçimsel kavramsallaştırılmalarının yazılmasına olanak sağlar. Esas gereksinimleri; iyi-tanımlı sözdizimi, iyi-tanımlı anlambilimi, etkili çıkarsama desteği, yeterli anlatımsal güç ve anlatım uygunluğudur [5]. İyi-tanımlı sözdiziminin önemi, bilginin makineler tarafından anlaşılabilmesi için gerekli koşulun olmasıdır. Dünya Çapında Ağ Birliği (World Wide Web Consortium - W3C) [65], dünya çapında ağ standartlarını belirleyen örgüttür. W3C; RDF, RDFS, OWL gibi anlamsal teknolojileri sunmaktadır. Alan ontolojilerinde kullanılan dillerden birisi olan OWL, aşağıdaki şekilde açıklanmıştır.

OWL: RDF diline bir eklentidir. OWL, sınıfların ve özelliklerin RDF anlamlarını kullanarak daha etkili açıklamaları destekleyen dil kurallarını içerir. OWL, aşağıdaki üç farklı alt dile ayrılmıştır [2].

- *OWL DL*: Sayısal yeterliliği yeniden elde etmek için kullanılır ve etkili çıkarsama desteği sağlar.
- *OWL Lite*: OWL DL kısıtlarını ve sınıflama hiyerarşisini destekler. Mantıksal sonuçları değerlendirecek sadece 0 veya 1 önemlilik değerini gönderir.
- *OWL Full*: Bütün OWL dillerinin özelliklerini içermektedir. OWL Full, RDF ile tam uyumlu olması sayesinde sözdizimsel ve anlamsal olarak tam uyumluluk sağlar.

### 3.1.2. Çoklu ortam Anlamsal Web Dilleri (Multimedia Semantic Web Languages)

Çoklu verinin anlamsal işlemleri, çoklu özelliklerin özne alan kavramları ile arasında bağlantının kurulması nedeniyle farklı bilgi gösterim şeklini gerektirir. Çoklu ortam içerik yapısının modellenmesi için M-OWL kullanılmaktadır. Varlıkları ve varlıkların hiyerarşik yapısını tanımlamak için M-OWL, OWL'daki benzer yapıları kullanmaktadır. Sınıflar ve sınıfların örnekleri, hiyerarşik yapıyı ve varlıkları tanımlamak için OWL temel elemanlarını kullanmaktadır [64].

Çoklu ortam ontoloji tanımlama dillerinden birisi de UzakEylem Nesnesi (TeleAction Object - TAO) 'dir. Bu dil, hiperçizge ve bilgi yapısı elemanları temelinde çoklu ortam nesnelere göstermek için kullanılan bir paradigmadır [66].

Videolarda gözlemlenen birçok kavram, bileşen varlıklarının belirli uzamsal-zamansal ilişkiler aracılığıyla tanımlanır. Çoklu ortam ontolojisi, bileşenler arasındaki uzamsal-zamansal ilişkiler açısından kavramları belirtir. M-OWL, belirli uzamsal-zamansal ilişkileri örnek ile desteklemek için kullanılan `<mowl:SpatialProperty>` ve `<mowl:Temporal Property>` olan iki alt sınıf tanımlı içermektedir. Bir alandaki belirsizlik, alan sınıfları ve sınıfların örnekleri üzerinde ortak olasılık dağılımı olarak tanımlanabilir.

### 3.1.3. Video Olay Ontolojileri Tanımlama Dilleri (Video Event Ontologies Definition Languages)

Videodaki bir olayın tanımlanması nesne, nesnelere arası ilişkiler ve ortam tanımlamalarını da kapsayacak şekilde olmalıdır. Literatürde bulunan bazı çalışmalarda, aşağıda örnekleri verilen Video Olay Gösterim Dili (Video Event Representation Language - VERL), Video Olay İşaretleme Dili (Video Event Markup Language - VEML) ve Olay Tanımlama Dili (Event Recognition Language - ERL) olay tanımlama dilleri önerilmektedir.

Nevatia vd. [67] çalışmalarında, video için olay tanımlama dili tanımlanmıştır. Videodaki olaylar, nesne hareketlerinin çeşitli zamansal ilişkileri şeklinde tanımlanmış daha basit olaylardan oluşan hiyerarşik bir yapıda olabilmektedir. Daha sonraki çalışmaları [68], video olay tanımlama için bir standartlaştırılmış taksonomi oluşturulma üzerinedir. Bu taksonomi, video etiketleme için VERL ve VEML'lerini kapsayacak şekilde tasarlanmıştır.

VERL: uygulama alan ontolojisinin tasarımı için ve ontolojideki kategoriler ile verinin etiketlenmesindeki olayların gösteriminde kullanılan dildir. VERL, olayları ve olayların birbiri ile etkileşimini tanımlar [68]. VERL'de, hiyerarşik çatıdaki düşük-seviye olaylar kullanılarak birleşik olaylar tanımlanabilir. VERL, adımlama ve değiştirme gibi daha karmaşık birleştirme işlemlerini içerir. Bileşik olaylar, çoklu-altüreçler olarak adlandırılan çoklu-eşzamanlı olayları içerir. Francois vd. [69] çalışmalarında, aralıklar arasındaki nicelik ilişkileri tanımlayan "Allen'in aralık cebri"ni kullanmışlardır.

VEML: video akışındaki olaylar ve nesnelere belirli örneklerini göstermek için kullanılmaktadır. VEML kullanılarak olay ismi, tipi, başlama ve bitiş zamanı, olaydaki aktörler XML tabanlı kodlanabilmektedir. VEML içinde bir sahne açıklaması, birkaç farklı tipteki veri akışını gerektirebilmektedir [68].

ERL: olay tanımlama ve sorgulama amaçları için varolan çeşitli karmaşıklıklarının olaylarını tanımlamak için kullanılır. ERL dört veri tipi içerir; nesne, yerleşim, aralık ve sayısal değer. Her veri tipi birçok özelliğe sahiptir. ERL'deki tanımlamalar,

programlama dillerindeki fonksiyon tanımlamaları ile benzerdir [67].

### 3.2. Çoklu ortam Ontolojileri (Multimedia Ontologies)

Genellikle çoklu ortam araştırmaları, işlenmemiş çoklu ortam içeriğinden otomatik olarak hesaplanmış düşük-seviyeli betimleyiciler ve yüksek-seviyeli özelliklerin çıkarımına odaklanmış anlamsallığı harmanlanması işi üzerinde yoğunlaşmışlardır [6]. Formalize edilmiş bilginin ve bir görüntünün görsel içeriğini tanımlayan özellikler kümesinin harmanlanması fikri; Voisine vd. [70] çalışmalarında verilen örnekte olduğu gibi, MPEG-7 [71] görsel özellikleri ve ontolojik bilgi kullanan bir bölge-tabanlı yaklaşım ile verilmektedir. Ayrıca, Snoek vd. [72] çalışmalarında, sözlük tarafından sürülen bir video etkileşimli bilgi elde etme yaklaşımını da önermişlerdir. Bazı araştırmalar “bag-of-words” yani “kelimeler çuvalı” yaklaşımı kategorisi altındadır. Buna göre, bir görüntü girdi görüntüsünün kesimlemesi veya kümelenmesi ile elde edilen “görsel kelimeler” kümesine ayrıştırılır.

Bir görüntü veya video belgelerinden elde edilmiş görsel özellikler iki kategoriye bölünmüştür. İlk kategori düşük-seviyeli özellikleri içerir. Bu kategori görsel özelliklerin nitel veya nicel tanımlarını sağlayabilmektedir. İkinci kategori yüksek-seviyeli özellikleri içermektedir. Bu kategori, bir görüntünün görsel içeriğini anlamsallık terimleri ile tanımlar [72].

Zengin anlam ile içerik tanımlarından otomatik olarak elde edilen yüzeysel anlam arasındaki anlamsal boşluğun doldurulması için ontolojiler yapısal video bilgileri ve görsel veri betimleyicileri içerecek şekilde zenginleştirilebilirler. Zengin bir biçimde video etiketleme daha karmaşık dilsel yapılar gerektirmektedir. Bu yapılar, anlamsal seviyede videoların bilgi gösterimi için kullanılırlar. Çoklu ortam ontolojileri, çoklu ortam bilgi gösterimini sağlamaktadır.

#### Çoklu ortam ontolojisinin özellikleri:

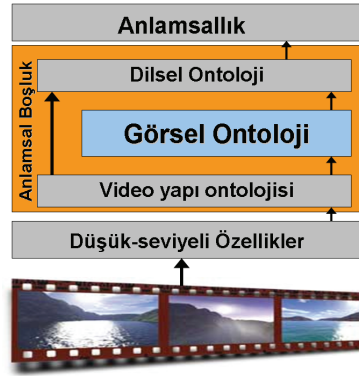
Bu ontolojiler, çoklu ortam bilgileri arasındaki ilişkilerinin tanımlanmasında, düşük-seviye niteliklerin yüksek-seviye niteliklerine eşlenmesini kolaylaştırmaktadırlar. Poli vd. [61] çalışmalarında çoklu ortam ontolojileri iki tipe ayrılmıştır. Birincisi ortama-özümlü ontolojiler, farklı ortam tiplerinin sınıflandırılmasını içerirler ve farklı ortam özelliklerini tanımlarlar. Diğer ise, içeriğe-özümlü ontolojiler, ortama özümlü olmamaları nedeniyle aynı alan ile ilgili diğer belgelerle kullanılabilirler.

#### Varolan çoklu ortam ontolojileri:

Varolan çoklu ortam ontolojileri, uygulama alanlarına veya çatılarına göre aşağıda verilen sınıflandırma gruplarına göre sınıflandırılmıştır;

- Çoklu ortam içerik yapı ontolojileri, çoklu ortam içerik yapısını tanımlamak için kullanılır.
- Çoklu ortam alan ontolojileri, çoklu ortam alanına ait ontolojilerin yaratılması için kullanılır.
- Çoklu ortam üst-seviye (upper) ontolojileri, yüksek-seviye kavramları tanımlar ve daha genel kullanım için tercih edilir.
- Çoklu ortam çekirdek ontolojileri, ontolojiler arasındaki köprü görevi gören ve farklı ontoloji tabanlı yaklaşımlar arasında karşılaştırmalar için referans noktası sağlamaktadır.
- MPEG-7 tabanlı çoklu ortam ontolojileri, MPEG-7 tanımlarını biçimlendirmek amacıyla kullanılır.

Şekil 2’de [73] anlamsal boşluğun doldurulması için ontolojilerin kullanımı şematik olarak gösterilmektedir.



Şekil 2. Anlamsal boşluğun doldurulması. (Filling the semantic gap)

Dilsel ontoloji: genel olarak veri hakkında veri tanımları yapılan üstveri olarak düşünülebilir [74]. Dilsel ontoloji, zaman uzayı, nedensellik ve insan etkileşim kavramlarındaki dilsel yapılar temelinde dilsel bilgi deposudur. Farrar vd. [75] çalışmalarında, dilsel ifadesi, birçok farklı alanlardaki kavramları kapsadığı için Standart Üst Birleştirilmiş Ontoloji (Standard Upper Merged Ontology - SUMO)’nin üstüne dilsel ontoloji yapılandırılması seçilmiştir. SUMO’nun seçilmesindeki neden, üst-seviye ontolojileri daha geniş bir kapsama alanı sunmak için harmanlamasıdır.

Görsel ontoloji: metinsel parçadan elde edilir ve ontolojinin her kavram ile ilişkili olan gösterimler kullanılarak zenginleştirilmiştir. Ayrıca görsel ontoloji, kavram hiyerarşisidir ve etiketleme kümelerine göre yapılandırılır [76].

Hollink vd. [77], 2005 yılındaki çalışmalarında görsel bilgi içeren bir ontolojinin geniş alanda etiketleme imkânlarını nasıl kolaylaştırabileceğini ve ontolojinin ihtiyacı olan gereksinimlerin nasıl tanımlanabileceğini incelemişlerdir. Bu istekleri temel alarak, görsel ve genel kavramlar arasında ilişki oluşturularak görsel bir ontoloji tanımlanmıştır.



Bimbo vd. [78], 2007 yılındaki çalışmalarında alan ontolojisi ile video yapı ontolojisi arasında bir köprü görevi yapan görsel prototiplerinin bulunduğu, video etiketleme ve arama için yeni bir ontoloji mimarisi denemişlerdir. MPEG-7 içinde çoklu ortam içeriği beş tip olarak; görüntü, video, ses, görsel-işitsel ve çoklu ortam olarak sınıflandırılmıştır [79].

Dasiopoulou vd. [80], 2009 yılındaki çalışmalarında aceMedia projesi kapsamında, Çoklu ortam Yapı Ontolojisi (Multimedia Structure Ontology - MSO) ve Görsel Betimleyici Ontolojisi (Visual Descriptor Ontology - VDO) olarak iki ontoloji geliştirmişlerdir. MSO, Çoklu ortam Betimleme Şemaları'ndaki (Multimedia Description Schemes - MDS) yapısal tanımlama araçlarının kümesini tamamlamaktadır. VDO ise görsel parçasını göstermektedir.

Görsel Betimleme Ontolojisi (VDO): Petridis vd. [81], 2006 yılındaki çalışmalarında çoklu ortam içerik açıklamadaki anlamsal boşluğu doldurmak için aceMedia projesi kapsamında ontoloji altyapısını uygulamışlardır. VDO, düşük-seviye görsel betimleyicilerle ontoloji kavramlarını bağlamak için geliştirilmiştir.

Aşağıda, çoklu ortam kavram ontolojileri verilmiştir.

Çoklu ortam için Büyük-ölçekli Kavram Ontolojisi (Large-Scale Concept Ontology for Multimedia - LSCOM) projesi, video etiketleme ve bilgi elde edilmesinde standart biçimsel sözlük tanımlama için oluşturulan bir çalıştırıdır. Naphade vd. [82], 2006 yılındaki çalışmalarında LSCOM sınıflandırmasını Cyc bilgi tabanına eşleme ile yapmışlardır. Eşleme, böylece Cyc altkütmesi ile ilişkili bir OWL dışa aktarımı üretecek ve OWL destekleyen LSCOM kavramları için Cyc içeriğine ait kısmı içerecektir.

LSCOM-Lite: LSCOM ontolojisinin ilk oluşturulan parçasıdır. TRECVID-2005 [44] ile uyumlu NIST tarafından yayınlanan otuzdokuz kavramdan oluşmaktadır.

### 3.2.1. Çoklu ortam İçerik Yapı Ontolojileri (Multimedia Content Structure Ontologies)

Bu tip ontolojiler çoklu ortam içerik yapısının tanımlanması ile ilgilidir. İçerik yapısı ontolojilerine örnek olarak aceMedia ve AIM@SHAPE projeleri verilebilir.

aceMedia (autonomous content entity Media) projesi, görsel-işitsel içerik analizi, nesne/olay tanımlama, bilgi elde etme, sahne tanımlama, zeki arama işlemlerini desteklemektedir [61].

AIM@SHAPE (Advanced and Innovative Models) projesinde, sayısal şekillerle ilişkili bilgi modelleme ve bilgi işleme için yeni bir yöntem geliştirilmektedir.

Bilgi; şekiller, yapılar, nitelikler, anlamsallık ve zaman etkileşimi ile ilgilidir [83].

### 3.2.2. Çoklu ortam Alan Ontolojileri (Multimedia Domain Ontologies)

Çoklu ortam alan ontolojileri, genellikle bir alanı modellemede kullanıldığı için çoklu ortam katmanlarından otomatik olarak elde edilen kavramları içermektedir. Bu ontolojiler iki role sahiptir. Birinci rolleri etiketleme yapılmasında kullanılan sözlükleri sağlamaktır. İkinci rolleri içeriğe bağlı ve kısıt çıkarsama ayırıştırma desteği için gerekli kavramsal ve uzamsal bilgi sağlamaktır. Örnek olarak Futbol, Formula 1 yarışları, Ortam Varlık ve Mücadele Ontolojisi (Media Presence and Campaign Ontology - MEPCO), Toplantı (Meeting), Yeni Bin Yıl için Yeni bir Ortam (The New Media for a New Millennium - NM2), SmartWeb Tümlüşik Ontolojisi (SmartWeb Integrated Ontology - SWIntO) gibi alan ontolojileri verilebilir.

Futbol alan ontolojisi: birçok futbol karşılaşmasının kaydedilmiş videolarını kullanarak futbol oyunu için alan ontolojileri de önerilmiştir [60, 73, 84]. Bu tarz alan ontolojiler, bu alana özgü futbol sahası ve futbol oyuncusuna ait görsel özellikleri içermektedir. Bilgi gösterim dili olarak OWL-DL kullanılmıştır.

Bertini vd. [60], 2007 yılındaki çalışmalarında, bir dilsel ontolojide tanımlanmış kavramların özelleştirilmesi olarak görsel prototiplerin eklendiği çoklu ortam ontolojilerini önermişlerdir. Bilgi gösterim dili olarak OWL kullanmışlardır.

Aldershoff vd. [85], 2003 yılındaki çalışmalarında, danışmanlı ve dinamik olan çoklu ortam nesne kategorizasyon sistemi önermişlerdir. Önerdikleri bu sistem, uzman insanların kategorizasyon davranışlarını öğrenerek, gerektiğinde futbol alan ontolojisindeki değişikliklere ve uzman görüşlerine göre kendini tekrar ayarlamaktadır.

Formula 1 alan ontolojisi: Formula 1 yarışlarının görsel-işitsel içeriğinin etiketlenmesi, bilgi elde etme ve filtreleme desteği için geliştirilmiştir. MPEG-7 üst-seviye ontolojisi, Formula 1 alanındaki bilgilerle genişletilmektedir. Bilgi gösterim dili olarak OWL-DL kullanılmıştır [84].

MEPCO: TV, basın, internetteki medya şirketleriyle ilişkisidir. MEPCO, PROTON üst-seviye ontolojisi temelindedir. Ortam ilişkili üstveri standartları (NewsML, News Codes gibi) ile karşılaştırılabilir [86].

Toplantı ontolojisi: toplantı zamanını ve yerini, katılımcıları, ses/video kayıtlarını, sunumlar gibi kaynakları ve gözden geçirme ile oluşturulan etiketlemeleri içermektedir. Bilgi gösterim dili olarak OWL kullanılmıştır [87].

**NM2 ontolojisi:** kural-tabanlı bileşeni ile birlikte seslendirme olayında kişiselleştirilmiş sürümler oluşturmak için geliştirilmiştir. Otomatik veya elle etiketleme, bilgi elde etme, içeriksel filtreleme ve çoklu ortam içerik eşleme alanında kullanılır. Kullanılan çoklu ortam alanı MPEG-7'dir [88].

**SWIntO:** Dilsel ve Bilişsel Mühendislik için Betimleyici Ontoloji (Descriptive Ontology for Linguistic and Cognitive Engineering - DOLCE), SUMO, MPEG-7 ontolojilerinin birleştirilmesi ile yapılandırılmış ontolojidir. Çoklu ortam verisini etiketleme ve elde etme işlemlerinde kullanılır [89].

### 3.2.3. Çoklu ortam Üst-Seviye Ontolojileri (Multimedia Upper Ontologies)

Üst-seviye ontolojileri, yüksek-seviye kavramları tanımlamakta ve daha genel amaçlı kullanılmasını sağlamaktadır. ZyX, MPEG-7 MDS, Anlamsal Görsel-İşitsel Eğlence Tekrar-kullanılabilen Nesnelere (Semantic Audiovisual Entertainment Reusable Objects - SALERO) örnek ontolojiler olarak verilebilir.

**ZyX üst-seviye ontolojisi,** çoklu ortam üst-modelleme için geliştirilmiş üst-seviye ontolojidir. Bilgi gösterim dili olarak OWL Full kullanılmıştır. ZyX çoklu ortam belge modeli belirtimleri temelindedir [88, 90].

**MPEG-7 MDS üst-seviye ontolojisi:** genel bir ontolojidir. Ontoloji bilgi gösterim dili olarak OWL DL kullanılmıştır. Tam olarak MPEG-7 MDS ve kısmen MPEG-7 görsel ve ses kısımlarını da içermektedir [84].

**SALERO:** ortam içeriğini tanımlamak için ontolojiler yaratmayı amaçlayan üst-seviye ontolojidir. Düşük-seviyeli anlamsal nitelikleri tanımlamak için MPEG-7 kullanılmaktadır. Yüksek-seviyeli anlamsal nitelikleri tanımlamak için ise Web Servis Modelleme Dili (Web Service Modeling Language - WSM) veya OWL, içerik bilgilerini tanımlamak için ise Anlamsal Web Kural Dili (Semantic Web Rule Language – SWRL) kullanılmaktadır [88].

### 3.2.4. Çoklu ortam Çekirdek Ontolojileri (Multimedia Core Ontologies)

Çoklu ortam çekirdek ontolojileri, alan bağımsız kavramların belirtimleri ve kavramlar arasındaki ilişkileri içeren belirtimlerin açık gösterimidir [61]. Çoklu ortam çekirdek ontolojilerine örnek olarak Uluslararası Dökümantasyon Komitesi'ne (International Committee for Documentation - CIDOC) ait Kavramsal Referans Modeli (Conceptual Reference Model - CRM) ve DOLCE verilebilir.

**CIDOC CRM:** çoklu ortam için çekirdek ontolojisinin parçası olarak kullanılmaktadır. CIDOC CRM, çoklu veri ve üstveri yapılarının anlamsallığı için sınıfların

yerine ilişki tanımlamaları üzerine yoğunlaşmıştır [61].

Doulaverakis vd. [91] çalışmalarında, heterojen dağıtık kültürel miras koleksiyonlarına birleşik erişimi sağlamak için bir yaklaşım önermişlerdir. Önerilen yaklaşım REACH Yunan Ulusal Projesi çatısı altında geliştirilmiştir. Bu proje; kültürel içeriğin ontolojik gösterimini, çoklu ortam tabanlı arama algoritmalarını, içerik-tabanlı erişim arayüzlerini içermektedir.

**DOLCE:** Bu ontoloji, bir üst-seviye ontolojisidir. Petridis vd. [92, 81] çalışmalarında farklı tipteki bölgeler arasındaki topolojik ve yönlü ilişkileri sağlamak için DOLCE'nin dolce:Region isimli bölge kavramını genişletmişlerdir. Yönlü uzamsal ilişkiler, görsel bölümlerin nasıl yerleştirileceğini tanımlar ve iki boyutlu veya üç boyutlu uzayda görsel bölümlerin birbirleriyle nasıl ilişkili olduklarını göstermektedir. Petridis vd. [93] çalışmalarında anlamsal çoklu ortam içerik analizi ve çıkarsama için bir bilgi gösterimi altyapısı; davranışsal modeller ve betimleyiciler düşük-seviye görsel-ışitsel özellikler içerecek şekilde otomatik içerik etiketlemenin yapılabilmesi için ontolojileri sunmaktadır.

### 3.2.5. MPEG-7 Tabanlı Çoklu ortam Ontolojileri (MPEG-7 Based Multimedia Ontologies)

MPEG-7 tanımlarını biçimlendirmek amacıyla ve çoklu ortam üstverisinin birlikte çalışabilirliğini artırmak için çok sayıda çoklu ortam ontolojileri önerilebilir. MPEG-7 tabanlı ontolojiler, daha yüksek seviyeli çoklu ortam tiplerini içerir. Sistem içerisinde çoklu ortam tiplerinin yapıları ve özellikleri de uygulanabilir [94]. Dasiopoulou vd. [80] 2009 yılındaki çalışmalarında, MPEG-7 tabanlı çoklu ortam ontolojileri gösterilmiştir. Bununla birlikte Troncy vd. [95] 2007 yılındaki çalışmalarında da dört tane MPEG-7 tabanlı çoklu ortam ontolojisi incelenmiştir. MPEG-7 tabanlı çoklu ortam ontolojilerine Hunter, aceMedia, SmartWeb, Çoklu ortam Bilgi Çıkarımı ile Önyüklemeye Ontolojisi Evrimi (Bootstrapping Ontology Evolution with Multimedia Information Extraction - BOEMIE), Alana Özgü-Çoklu ortam Endeksleme-Bilgi elde etme-Filtreleme (Domain Specific-Multimedia Indexing-Retrieval-Filtering - DS-MIRF), Rhizomik, Çoklu ortam Çekirdek Ontolojisi (Core Ontology of MultiMedia - COMM) ontolojileri örnek verilebilir.

**Hunter ontolojisi:** Harmony projesi kapsamında Hunter tarafından önerildiği için Hunter MPEG-7 ontolojisi [95] genellikle, büyük anlamsal çatıları kullanan ontolojinin sanal betimleyicileri ve görüntülerin ayrışmalarını tanımlamak için uygulanmaktadır. ABC üstveri modeli, farklı alandaki üstveri ontolojileri arasındaki birlikte çalışabilirliği kolaylaştırmaktadır [61].

**aceMedia ontolojisi:** aceMedia projesi kapsamında MSO ve VDO olmak üzere iki RDFS ontolojisi geliştirilmiştir. Modelleme için MSO'nun ve VDO'nun geliştirilmesinde Hunter'ın ontolojisine benzer bir sınıf hiyerarşisi ile sonuçlanan aynı mühendislik prensiplerini takip eden bir model seçilmiştir [88].

**SmartWeb ontolojisi:** bu ontolojide Çoklu ortam İçerik sınıfları, ilgili Kesim sınıfları olarak görülmez. Ayrıca, farklı bir modelleme perspektifinin uyarlanması ile ayrışım şema sınıfları olarak modellenmiş, her biri içerik/bölüm başına geçerli bir ayrışım örüntüsü ve bir uzamsal-zamansal boyutu ifade etmektedir [80].

**BOEMIE ontolojisi:** bu ontolojinin geliştirildiği çoklu ortam içerik analizi projesi kapsamında, MPEG-7 MDS yapısal tanımların anlamsallığını elde etmek için Çoklu ortam İçerik Ontolojisi (Multimedia Content Ontology - MCO) ve Çoklu ortam Betimleyici Ontolojisi (Multimedia Descriptors Ontology - MDO) olarak iki OWL-DL ontolojisi geliştirilmiştir. MDO düşük-seviye özelliklerin gösterimini destekler. MCO, çıkarsamaya yüksek-seviye anlamsal yorumlamalar kazandırabilmek için sınıflandırma sonuçlarının gösterimini destekler [80].

**DS-MIRF ontolojisi,** MPEG-7 MDS ve Sınıflandırma Şemalarını (SS) anlamsal olarak ifade edebilmek için OWL-DL tabanlıdır. Alan bilgisinin MPEG-7 yapısına sistematik şekilde tümlleştirilmesi için Futbol ve Formula 1 OWL alan ontolojileri ile bütünleşmektedir [84].

**Rhizomik ontolojisi:** bütün MPEG-7 standartının, OWL ontoloji gösterim diline tam ve otomatik eşlemesi sonucunda, çoklu ortam üstverisi için üst-seviye ontoloji olarak MPEG-7 OWL ontolojisi üretilmiştir. [80, 96].

**COMM:** çoklu ortam örüntülerinden oluşmaktadır. OWL-DL tabanlıdır. Varolan web ontolojileri ile birlikte çalışabilirlik özelliği sayesinde çoklu ortam etiketlemesi için MPEG-7'den daha fazla imkâna sahiptir [97].

Tablo 4'de MPEG-7 tabanlı çoklu ortam ontolojilerinin kaynaklar, karmaşıklık, kapsama alanı, kaynakça ve uygulamalar olarak beş ana özellik özetlenmiştir. Bu çalışma kapsamında incelenen çoklu ortam ontolojilerinin şematik olarak gösterimi Şekil 3'de verilmiştir.

Garcia ve Celma [96], 2005 yılındaki çalışmalarında geliştirdikleri yaklaşım tüm MPEG-7 standartını OWL'a tam biçimde eşleme amaçlıdır. Bu yaklaşım, genel bir XML Şema'yı OWL ontoloji gösterim diline eşleme temellidir. Bu eşleme sayesinde çalışmaya ait MPEG-7 OWL ontolojisini oluşturmuşlardır. Bu ontoloji, 2372 sınıf ve 975 özellikten oluşmaktadır.

Tsinaraki vd. [98], 2004 yılındaki çalışmalarında OWL ile tanımlanmış alan ontolojileri, MPEG-7 MDS üstveri ile tümlleştirilebilir hale getirilmişlerdir. Bu çalışmada, alan ontolojileri için endeksleme, bilgi elde etme, filtreleme işlemleri MPEG-7 MDS yapısı ile tanımlanabilir hale gelmiştir. Önerdikleri yöntem, bir OWL üst-seviye ontolojidir.

**Tablo 4.** MPEG-7 tabanlı çoklu ortam ontolojileri (Multimedia ontologies based on MPEG-7).

|                      | Harmony Ontolojisi     | aceMedia Ontolojisi               | SmartWeb Ontolojisi               | BOEMIE Ontolojisi                 | DS-MIRF Ontolojisi       | Rhizomik Ontolojisi                                    | COMM Ontolojisi                   |
|----------------------|------------------------|-----------------------------------|-----------------------------------|-----------------------------------|--------------------------|--------------------------------------------------------|-----------------------------------|
| <b>Kaynaklar</b>     | ABC                    | DOLCE                             | DOLCE, SUMO                       | Yok                               | Yok                      | Yok                                                    | DOLCE                             |
| <b>Karmaşıklık</b>   | OWL Full               | RDFS                              | OWL, RDFS                         | OWL-DL                            | OWL-DL                   | OWL-DL                                                 | OWL-DL                            |
| <b>Kapsama alanı</b> | MDS+ Görsel            | MDS+ Görsel                       | MDS+ Görsel                       | MDS+ Görsel-işitsel               | MDS+SS                   | Hepsi                                                  | MDS+ Görsel                       |
| <b>Kaynakça</b>      | Hunter vd., 2001 [107] | ACEMedia projesi, 2009 [88]       | Dasiopoulou vd., 2009 [80]        | Dasiopoulou vd., 2009 [80]        | Tsinaraki vd., 2007 [84] | Dasiopoulou vd., 2009 [80], Garcia ve Celma, 2005 [96] | Arndt vd., 2007 [97]              |
| <b>Uygulamaları</b>  | Sayısal kütüphaneler   | Çoklu ortam Analizi ve Etiketleme | Çoklu ortam Analizi ve Etiketleme | Çoklu ortam Analizi ve Etiketleme | Sayısal kütüphaneler     | Sayısal Hakların yönetimi, e-iş                        | Çoklu ortam Analizi ve Etiketleme |



**Şekil 3.** Çalışmada kullanılan çoklu ortam ontolojilerin hiyerarşisi (The hierarchy of multimedia ontologies used in our study).

Bertini vd. [99], 2008 yılındaki çalışmalarında hem dilsel hem de görsel ontoloji içeren bir çoklu ortam ontolojisini tanımlamışlardır. Bu ontoloji, basit olayları, özneleri veya nesnelere tespit için kullanılan uygun düşük-seviye betimleyicileri içermektedir.

#### Video endeksleme:

Lee vd. [100], 2008 yılındaki çalışmalarında ontoloji tabanlı video endeksleme için nesnelere tespiti ve etkin bir ontolojinin nasıl yapılacağı konuları açıklanmıştır. Tüm nesnelere tanımlanmıştır ve ontoloji işleme vasıtasıyla içerikleri elde edilmiştir. Bu yaklaşım ile video endekslemede tek bir ontoloji oluşturulmuş ve sunulmuştur.

Mezaris vd. [101], 2004 yılındaki çalışmalarında nesne-tabanlı video endeksleme ve bilgi elde etme için yeni bir yaklaşım önermişlerdir. Gerçek zamanlı, sıkıştırılmış görüntü dizilerinin danışmansız kesimlenmesi ve bilgi elde etmek için ontolojileri içermektedir.

Jiang vd. [102], 2008 yılındaki çalışmalarında yeni bir Bag-of-visual-Words (Görsel kelimeler çuvalı - BoW) genişleme yöntemi önermişlerdir. Bununla görsel kelime düzeltme probleminin etkisini hafifletmişlerdir. Bu iş görsel kelimelerin bağlılığını temel alan BoW'deki görsel kelimelerin ağırlıklarının yayılmasıyla yapılmıştır.

Benmokhtar vd. [103] 2009 yılındaki çalışmalarında, daha önceki çalışmalarındaki içerik tabanlı video çekimi endekslemeyi TRECVID-2007 videodan bilgi elde etme işlemi için ontoloji tabanlı kavram yapısını gözönüne alarak iyileştirmişlerdir ve ontoloji tabanlı bir video endeksleme sistemi geliştirmişlerdir. Bu çalışma iki adımdan oluşmaktadır. İlkinde her bir tek kavram bağımsız olarak modellenmiştir. İkincisinde bir ontoloji tabanlı kavram güven değerlerinin ontolojik tekrar ayarlanması ve kavramlar arasındaki akıcı ilişkilerinin gösterimi yoluyla sunulmuştur.

#### Video etiketleme ve özetleme:

Park vd. [10] çalışmalarında, OLYVIA isminde yeni bir otomatik video etiketleme ve özetleme sistemi önermişlerdir. Bu sistem, video bilgi elde etmeyi kolaylaştırıcı anlamsal çıkarsama kuralları ve ontolojileri içermektedir. Çalışmalarında, video çekimi/grup/sahne/video seviyesine ait yüksek-seviyeli kavramlar, VideoAnnotation ontolojisi ve Object ontolojisi için anlamsal çıkarsama kurallarının uygulanması ile otomatik olarak elde edilmişlerdir.

Bagdanov vd. [73], 2007 yılındaki çalışmalarında hem dilsel hemde dinamik görsel ontolojilerin her ikisini de içeren çoklu ortam ontolojileri verilmiştir. Bu çalışmada açıklanan çoklu ortam ontolojileri, görsel ve dilsel kavramları görsel prototipleme yoluyla ilişkilendirilmesine dayanır. Görsel prototipler düşük-seviye kavramların ve video yapı ontolojisinde tanımlanan kavram betimleyicilerinin örnekleridir.

Bertini vd. [104], 2008 yılındaki bir diğer çalışmalarında anlamsal video etiketleme yapan ontolojileri kullanan algoritmalar ve güncel yaklaşımları ele alarak yüksek seviyeli kavramları tanımlayan kuralları otomatik olarak öğrenen bir yaklaşım sunmuşlardır. Bu yaklaşım, anlamsal video etiketlemede kurallar kümesini öğrenmek için bir ontoloji içerisine gömülmüş alan bilgisini kullanır.

Ballan vd. [105], 2009 yılındaki çalışmalarında futbol videolarının anlamsal etiketlenmesi için Dinamik Resimsel Zenginleştirilmiş Ontoloji (Dynamic Pictorial Enriched Ontologies - DPEO) olarak tarif edilen bir ontoloji modelini kullanan bir çatıyı önermişlerdir. Etiketleme için üç adet mekanizma desteklenmektedir; görsel örnek kümeleme, prototip seçimi ve dinamik küme güncelleme. Ontolojideki görsel örnekler, benzer hareket veya görsel örüntülerin kümelerine ait görsel prototipler şeklinde ontolojide içerilmektedir.

#### Video olay analizi:

Song vd. [106], 2005 yılındaki çalışmalarında video olay analizi ve alan bilgi ontolojisi tabanlı tanımlama için yeni bir yöntem önermişlerdir. Video olaylarının içeriğindeki anlamsal kavramlar, alan bağımsız faktörler (piksel rengi, hareket vektörleri ve uzay-zamansal ilişki) ve anlamsal nesnelere nitel özellikleri ile zenginleştirilmiş bir özel alanda tanımlanmıştır. Bu sayede düşük-seviye özellikler ve yüksek-seviye anlamlar arasındaki boşluğu gidermişlerdir.

Nevatia vd. [67], 2003 yılındaki çalışmalarında fiziksel dünyadaki uzamsal-zamansal karmaşık olayların basit olayların bir kompozisyonu şeklinde gösterilebildiği bir olay ontoloji tanımlamışlardır. Basitten daha karmaşığa doğru olay kompozisyonu, olayların bir dil tanımlamasını da ortaya çıkarmıştır. Kullanıcılar, ERL kullanılarak olay tanımlama ve sorgulama işlemlerini yapabilirler.

SanMiguel vd. [57] 2009 çalışmalarında, video olay analizi ile ilgili önceki bilginin gösterimi için bir ontoloji önermişlerdir. Bu ontoloji uygulama alanı ve analiz sistemi ile ilişkili olan alan bilgisi ve sistem bilgisinden oluşmaktadır. Önerilen ontoloji, temel ontoloji ve alana özgü genişlemelerden oluşmaktadır.

### **3.3. Çoklu ortam İçerik Tanımlama Standartları ve Protokolleri (Multimedia Content Definition Standards and Protocols)**

MPEG-7 standartının hedefi, görsel-ışitsel tanımları oluşturma ve insanların ve makinelerin birlikte çalışabilirliğine olanak sağlayan standartlaştırılmış araç geliştirmektir. MPEG-7 çeşitli ortam tiplerini (görüntüler, grafikler, üç boyutlu modeller, ses, konuşma, video ve bunların bileşimleri) ele almaktadır. MPEG-7 tanımlama şemaları ve betimleyicileri sadece XML Şema ile ifade edilir.

MPEG-7 standartının ana elemanları aşağıda verilmiştir [71].

- *Betimleyiciler (Descriptors - D)*, her özellik gösteriminin sözdizimsel ve anlambilimini (semantics) tanımlar.
- *Betitleme Şemaları (Description Schemes - DS)*, DS bileşenleri arasındaki ilişkilerin anlambilimini ve yapısını tanımlar.
- *Betitleme Tanım Dili (Description Definition Language - DDL)*, yeni DS'nin yaratılmasına ve D'nin yaratılmasına ve varolan DS'nin değiştirilmesine ve genişletilmesine de izin verir.
- *Sistem araçları (System tools)*, içerik ile betimlemelerinin senkronizasyonunu, depolama ve iletim için iletim mekanizmaları ve kodlanmış gösterimleri, MPEG-7 betimlemelerindeki zihinsel özelliğin korunmasını destekler.

Hunter [107], 2001 yılındaki çalışmasında MPEG-7 için RDFS içerisinde gösterilen bir ontoloji geliştirmiştir ve MPEG-7 için RDFS anlamsal tanımlarının nasıl betimlendiğini göstermiştir. MPEG-7'de çoklu ortam içeriği görüntü, video, ses, görsel-işitsel ve metin olarak beş tipte sınıflandırılmıştır. Her tip, kendi segment alt sınıflarına sahiptir.

Çoklu ortam üstverisinin sözdizimsel birlikte çalışabilirliğinin en basit tipi belirli sayıda standart ve protokolün kullanımı ile başarılı olmaktadır. Bu standart ve protokoller arasında MPEG-7, MPEG-21, CIDOC CRM, Dublin Core bulunmaktadır. MPEG-21'de, içerik yaratıcıları, üreticiler, dağıtıcılar ve servis sağlayıcılarını destekler. Dublin Core, web ortamında bilgi kaynaklarının tanımlanması için bir üstveri standarttır. Çoklu ortam içerik tanımlama standartları ve protokolleri bir sözdizimini tanımlamak için yüksek-seviyeli bir model veya XML gibi özel bir şema kullanırlar. Böyle şemaların uygulamaları DELOS projesi gibi projelerdeki protokoller ise Açık Arşiv İniyatifi (Open Archives Initiative - OAI) [108], TV-Anytime [109, 110], NewsML [111, 112] gibi kullanılır. TV-Anytime [110], sadece anahtar kelime tabanlı anlamsal tanımlama yeteneği sağlar. NewsML [112], XML tabanlı haberlerin gösterimi ve yönetiminde bir seviye yüksek standarttır. OAI, bilgi toplama için sayısal ortam gösterimlerine ait tanımları oluşturan bir protokolü yasaklar.

### 3.4. MPEG-7 Etiketleme Araçları (MPEG-7 Annotation Tools)

MPEG-7 standartının temeli, düşük-seviyeli görsel özelliklerin çıkarılması yoluyla görüntü ve video verisi endekslemektir [71]. Buna en güzel örnek olarak Otomatik İçerik Varlığı (Autonomous Content Entity - ACE) yazılımsal alet kutusu [113] verilebilir. Bir içerik ön işleme yazılımsal alet kutusu olan aceToolbox, aceMedia projesi dâhilinde geliştirilmiş olup, düşük-seviyeli analiz ve MPEG-7 özellik çıkarımından sorumludur. Başka bir örnek ise, halka

açık bir şekilde sunulan M-Onto-Mat-Annotizer [92, 81] isimli etiketleme yazılımsal aracıdır. M-Onto-Mat-Annotizer görsel düzenleyici ve ortam gösterici, görsel içeriğin işlenmesi ve yüklenmesi, görsel özelliklerin çıkarımı ve alan ontolojileri kavramları ile ilişkilendirilmeleri için grafiksel bir arayüze sahiptir.

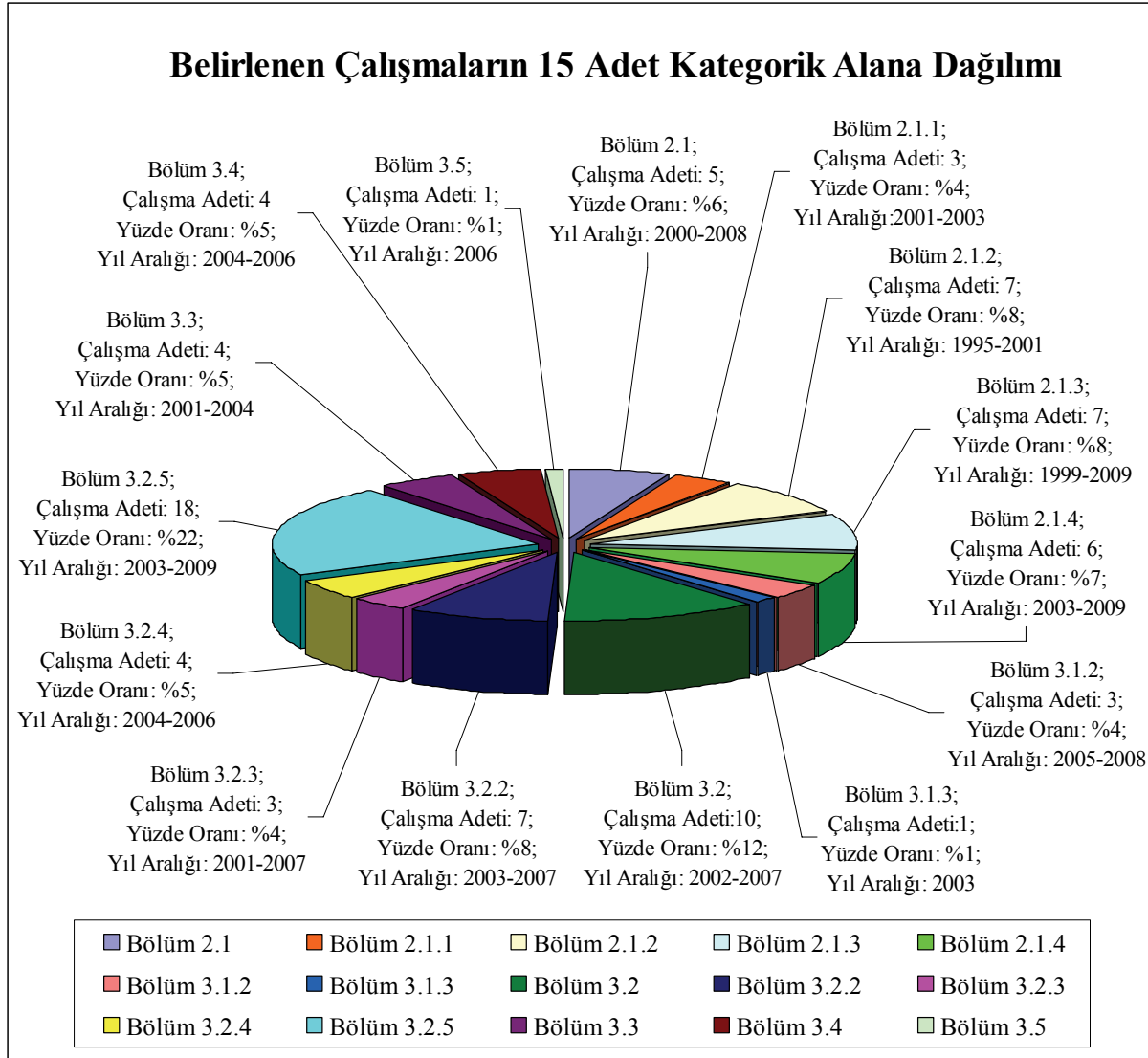
Protégé [114], bir kullanıcının OWL ontolojilerini oluşturma, yükleme, veriyi etiketleme ve etiket işaretlerini kaydetme işlemlerine olanak tanımaktadır. Bununla beraber, Ortam Slot Unsuru (Media Slot Widget - MSW) sayesinde sadece basit çoklu ortam desteği sağlamaktadır. Ortam Slot Unsuru, ses ve video dosyalarının gösterilmesini ve içerilmesini sağlar.

IBM MPEG-7 Etiketleme aracı [115], IBM tarafından video çekimlerini temel alan MPEG-7 üstverisi ile video dizilerini etiketlemeye yardımcı olması için geliştirilmiştir, fakat herhangi bir ontoloji dilini desteklemez. Bu araç, bir kullanıcının video çekimlerini etiketlemek için seçebildiği anahtar kelimeleri içeren düzenlenebilir bir sözlük kullanır.

## 4. SONUÇLAR VE TARTIŞMA (RESULTS AND DISCUSSIONS)

Çoklu ortam analizi günümüzde oldukça popüler bir araştırma alanıdır. Anlamsal Web ile çoklu ortam belgelerinin belirli bir sorgu, endeksleme, özetleme vb. işlemleri yoluyla etkin kullanımı ve sınıflandırılması, karmaşık ilişkilerin organize edilmesi ve zaman kaybının önüne geçilmesi için vazgeçilmez bir hale gelmektedir. Çalışmamızda, varolan ontolojiler ve çoklu ortam ontolojileri başta olmak üzere, bunlara temel oluşturan diğer konularda incelenmiştir.

Literatür taraması sonucu, 1995 ile 2009 yılları arasında yapılan çeşitli çalışmalar 15 adet kategorik alana ayrıştırılmıştır. Tablo 5'de hangi yıllar arasında, hangi alanda daha fazla çalışma yapıldığı ve bu alanlarda kaç adet yayın çıkarıldığı özetlenmektedir. İncelenen herhangi bir çalışmanın içerdiği teknik özellikler göz önüne alınarak; uygun bulunmuş ise birden çok kategoriye de dâhil edilmiştir. Örneğin Bertini vd. [60] çalışmalarında, hem çoklu ortam alan ontolojileri konusunu hem de video veri seti içeren başarımlar ölçüm ve değerlendirme konusunu kendi içeriklerinde sunmaktadırlar. Çalışmamızda belirlenen kategoriler de yer vermediğimiz daha genel çaptaki 31 adet çalışma (kitaplar, inceleme yazıları, teknik raporlar vs.) ve 16 adet web sitesi de makale içerisinde bulunmaktadır. Bunlar literatür taramasındaki ilgili alanlar için belirleyici nitelikte olmadıklarından yüzde oranlarına katılmamışlardır. Tablo 5'de bu kategorik alanlar makalemizde geçtiği şekilde ait oldukları bölüm numaraları ile belirtilmektedirler. Ayrıca bu tablodaki verilere göre Şekil 4'da bu kategorilere ait grafiksel gösterimler verilmiştir.



**Şekil 4.** Belirlenen çalışmaların 15 adet kategorik alana dağılımı (Distribution of determined studies for 15 categorical areas)

**Tablo 5.** Literatür taraması ile belirlenen çalışmaların 15 adet kategorik alana dağılımı. (Distribution of studies for 15 categorical area determined in survey)

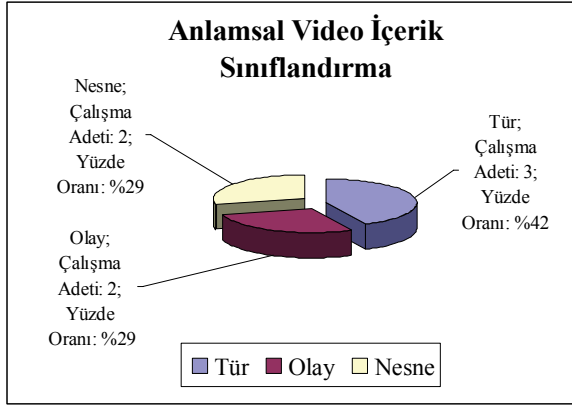
| Kategori No | Bölüm No    | Çalışma Adeti | Yüzde Oranı | Yıl Aralığı |
|-------------|-------------|---------------|-------------|-------------|
| 1           | Bölüm 2.1   | 5             | %6,0240     | 2000-2008   |
| 2           | Bölüm 2.1.1 | 3             | %3,6144     | 2001-2003   |
| 3           | Bölüm 2.1.2 | 7             | %8,4336     | 1995-2001   |
| 4           | Bölüm 2.1.3 | 7             | %8,4336     | 1999-2009   |
| 5           | Bölüm 2.1.4 | 6             | %7,2288     | 2003-2009   |
| 6           | Bölüm 3.1.2 | 3             | %3,6144     | 2005-2008   |
| 7           | Bölüm 3.1.3 | 1             | %1,2048     | 2003        |
| 8           | Bölüm 3.2   | 10            | %12,0480    | 2002-2007   |
| 9           | Bölüm 3.2.2 | 7             | %8,4336     | 2003-2007   |
| 10          | Bölüm 3.2.3 | 3             | %3,6144     | 2001-2007   |
| 11          | Bölüm 3.2.4 | 4             | %4,8192     | 2004-2006   |
| 12          | Bölüm 3.2.5 | 18            | %21,6864    | 2003-2009   |
| 13          | Bölüm 3.3   | 4             | %4,8192     | 2001-2004   |
| 14          | Bölüm 3.4   | 4             | %4,8192     | 2004-2006   |
| 15          | Bölüm 3.5   | 1             | %1,2048     | 2006        |

Anlamsal video içerik sınıflandırma işlemi tür (cins), olay ve nesne olarak üç farklı şekilde ele alınmıştır. Bu inceleme sonucunda, literatürdeki bu çalışmaların yıl ve oran bazındaki dağılımları Tablo 6'da verilmektedir. Tablo 6'da makalemize ait bölümler ait oldukları bölüm numarası ile verilmektedir.

**Tablo 6.** Anlamsal video içerik sınıflandırmanın çalışmalara dağılım oranları. (Distribution rate of semantic video content classification for studies)

| Anlamsal Video İçerik Sınıflandırma | Çalışma Adeti | Yüzde Oranı |
|-------------------------------------|---------------|-------------|
| Tür                                 | 3             | %42,8571    |
| Olay                                | 2             | %28,5714    |
| Nesne                               | 2             | %28,5714    |

Anlamsal video içerik sınıflandırma işlemi ile ilgili olan literatürdeki çalışmaların yıl ve oran bazında dağılımları Şekil 5'de grafiksel olarak gösterilmektedir.



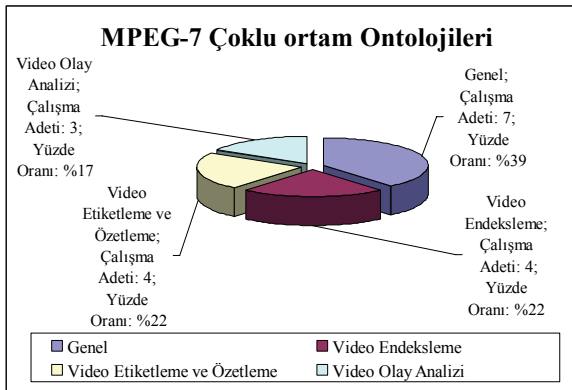
Şekil 5. Anlamsal video içerik sınıflandırma. (Semantic video content classification)

MPEG-7 tabanlı çoklu ortam ontolojileri çalışmamızda incelenmiştir. Bu inceleme sonucunda, literatürdeki çalışmalar yıl ve oran bazındaki dağılımları Tablo 7’de verilmektedir. Tablo 7’de makalemize ait bölümler ait oldukları bölüm numarası ile geçmektedir.

Tablo 7. MPEG-7 Çoklu ortam ontolojilerinin çalışmalara dağılım oranları. (Distribution rates of MPEG-7 multimedia ontologies for studies)

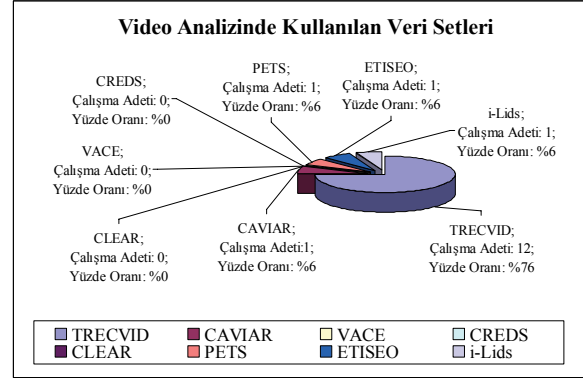
| MPEG-7 Çoklu ortam Ontolojileri | Çalışma Adeti | Yüzde Oranı |
|---------------------------------|---------------|-------------|
| Genel                           | 7             | %38,8888    |
| Video Endekleme                 | 4             | %22,2222    |
| Video Etiketleme ve Özetleme    | 4             | %22,2222    |
| Video Olay Analizi              | 3             | %16,6666    |

MPEG-7 Çoklu ortam ontolojilerinin literatürdeki çalışmalarda kullanım oranları Şekil 6’de grafiksel olarak gösterilmektedir.



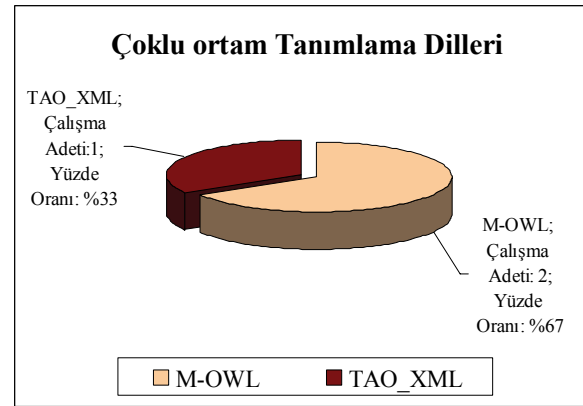
Şekil 6. MPEG-7 Çoklu ortam ontolojileri. (MPEG-7 Multimedia ontologies)

Çalışmamızda ele alınan video analizinde kullanılan veri setleri: TRECVID, CAVIAR, VACE, CREDs, CLEAR, PETS, ETISEO ve i-Lids’dir. Bunlar incelendiğinde, literatürdeki çalışmalarda kullanım oranları Şekil 7’de grafiksel olarak gösterilmektedir.



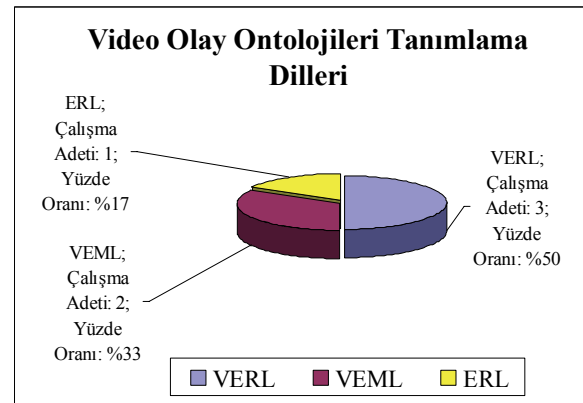
Şekil 7. Video analizinde kullanılan veri setleri. (Datasets used for video analysis)

Çoklu ortam Anlamsal Web dilleri: M-OWL ve TAO\_XML’dir. Bunlar incelendiğinde, literatürdeki çalışmalarda kullanım oranları Şekil 8’de grafiksel olarak gösterilmektedir.



Şekil 8. Çoklu ortam tanımlama dilleri. (Multimedia definition languages)

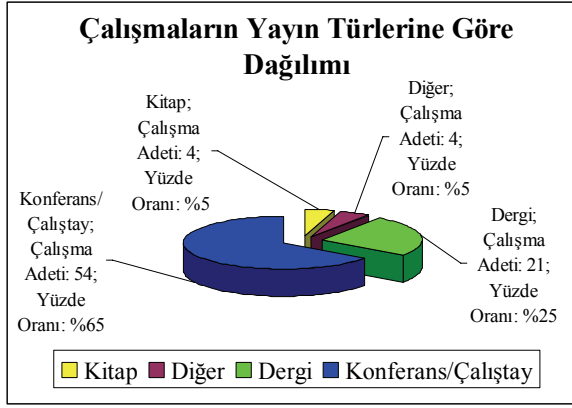
Video olay ontolojileri tanımlama dilleri: VERL, VEML ve ERL’dir. Bunlar incelendiğinde, literatürdeki çalışmalarda kullanım oranları Şekil 9’de grafiksel olarak gösterilmektedir.



Şekil 9. Video olay ontolojileri tanımlama dilleri. (Video event ontologies definition languages)

Belirlediğimiz 15 adet alan için kategorize edilmiş çalışmalara ait yayın türleri Şekil 10’de grafiksel olarak gösterilmektedir.





**Şekil 10.** Çalışmaların yayın türlerine göre dağılımı  
(Distribution of studies according to publication types)

Literatürdeki çalışmaların gösterdiği güncel durum, anlamsal çoklu ortam/video analizi'nin her aşama ve tür (cins) için anlamsal çıkarsamaya temel oluşturduğudur. Bu temel; bir analiz yapılması ile yüksek başarımlı sağlayabilmektedir. MPEG-7 ve MPEG-21 gibi anlamsal içerik kullanımını destekleyen standartlar ve protokoller, ontoloji kullanımının yaygınlaşmasına neden olmuşlardır. Video olay tanımlama ontolojileri ve bunlara ait diller sayesinde, video olayları gibi karmaşık durumların tarif edilebilmesi ve makine tarafından anlaşılabilir bir biçime dönüşmesi daha kolay hale gelmiştir. Bu sayede analiz hızlanmakta ve elde edilen sonuçlar daha tutarlı hale gelmektedir. Çoklu ortam analizi ve onun özelinde videoların anlamsal olay analizi, önümüzdeki günlerde de popülerliğini artırarak konu ve çalışma bazında genişlemeye devam edecektir.

#### KAYNAKLAR (REFERENCES)

1. Farag, W. E., ve Abdel-Wahab, H., "Adaptive Summarization of Digital Video Data", **Video Data Management and Information Retrieval**, Editör: S. Deb, Hershey, PA: Idea Group Publishing, A.B.D., 77-99, 2005.
2. Daconta, M., C., Obrst, L., J., ve Smith, K., T., **The Semantic Web, a Guide to the Future of XML, Web Services and Knowledge Management**, 85-116, Wiley, A.B.D. 2003.
3. Ossenbruggen, J., V., Stamou, G, ve Pan, Z., J., "Multimedia Annotations and the Semantic Web", **SWCASE05**, Semantic Web Case Studies and Best Practices for eBusiness, November 7, 2005.
4. Grosky, W., I., Sreenath, D., V., ve Fotouhi, F., "Emergent Semantics and the Multimedia Semantic Web", **SIGMOD Rec.**, Cilt. 31, No. 4, 54-58, ACM Press, 2002.
5. Garcia, R., Tsinaraki, C., Celma, O., ve Christodoulakis, S., "Multimedia Content Description using Semantic Web Languages", **Semantic multimedia and ontologies: theory and applications**, 17-54, Bölüm 2, Springer-Verlag, ISBN: 978-1-84800-075-9, 2007.

6. Seremeti, L., ve Kameas, A., "Multimedia ontologies", **MobiMedia '07: Proceedings of the 3rd international conference on Mobile multimedia communications, Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering - ICST**, Brüksel, Belçika, 1-7, 2007.
7. Bertini, M., Cucchiara, R., Bimbo, A.,D., ve Torniai, C., "Ontologies Enriched with visual Information for Video Annotation", **European Semantic Web Conference -ESWC05**, 2005.
8. Ghosh, H., Chaudhury, S., Kashyap, K., ve Maiti, B., "Ontology Specification and Integration for Multimedia Applications", **Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems**, Bölüm 9, Springer, 2007.
9. Ding, L., Kolari, P., Ding, Z., ve Avancha, S., "Using Ontologies in the Semantic Web: A Survey", **Ontologies: A Handbook of Principles, Concepts and Applications in Information Systems**, Editör: R. Sharman, R. Kishor ve R. Ramesh, Bölüm 4, Springer, 79-114, 2006.
10. Park, K-W, Jeong, H-L, Moon, Y. S., Park, S. H. ve Lee, D-H, "OLYVIA: Ontology-based Automatic Video Annotation and Summarization System Using Semantic Inference Rules", **Third International Conference on Semantics, Knowledge and Grid - SKG 2007**, 170-175, Xian, China, 2007.
11. Hanjalic, A., "Towards Theoretical Performance Limits of Video Parsing", **IEEE Transactions on Circuits and Systems for Video Technology**, Cilt 17, No.3, 261-272, 2007.
12. Hongeng, S. ve Nevatia, R., "Multi-agent event recognition", **Proceedings of International Conference Computer Vision**, 84-93, 2001.
13. Porikli, F., **Video Object Segmentation**, Doktora tezi, Polytechnic University, (UMI dissertation publishing), Brooklyn, Newyork, A.B.D., 2001.
14. Farag, W. E., ve Abdel-Wahab, H., "Video Database Techniques and Video-on-demand", **Multimedia Systems and Content-based Image Retrieval**, Editör S. Deb, Hershey, PA: Idea Group Publishing, A.B.D., 114-154, 2004.
15. Dimitrovski, I., Lokovska, S., Kakasevski, G, ve Chobev, I., "Video Content-Based Retrieval System", **The International Conference on "Computer as a Tool" -EUROCON 2007**, Varşova, Polonya, 978-983, 2007.
16. Dönderler, M. E., Ulusoy, O., ve Güdükbay, U., "A rule-based video database system architecture", **Information Sciences**, Cilt. 143, 13-45, 2002.
17. Köprülü, M., Çicekli, N.K., ve Yazıcı, A., "Spatio-temporal querying in video databases", **Information Sciences**, Cilt 160, 131-152, 2004.
18. Kuo, T.C.T. ve Chen, A.L.P., "Content-based query processing for video databases", **IEEE**



- Transactions on Multimedia**, Cilt 1, No. 2, 1-13, 2000.
19. Acharya, B., Majumdar, A. K., ve Mukherjee, J., "Video model for dynamic objects", **Information Sciences**, Cilt. 176, 2567-2602, 2006.
  20. Jung, M. Y. ve Park, S. H., "Semantic-Based Scene Retrieval Using Ontologies for Video Server", **Proceeding of the 23rd International Technical Conference on Circuits/Systems, Computers and Communications, ITC-CSCC2008**, 45-48, 2008.
  21. Hauptmann, A. Ng, T.D. Baron, R., Lin, W., M., Chen, Derthick, M., Christel, M., Jin, R., ve Yan, R., "Video classification and retrieval with the Informedia digital video library system", **Text Retrieval Conference - TREC02**, 2002.
  22. Shahraray B., "Multimedia information retrieval using pictorial transcripts", **Handbook of Multimedia Computing**, Editör: B. Furth, CRC Press, Boca Raton, FL, A.B.D. 345-359, 1999.
  23. Sebe N., Lew, M. S., ve Smeulders, A. W. M., "Video retrieval and summarization", **Computer Vision and Image Understanding, Special Issue on Video Retrieval and Summarization**, Editorial Introduction, Cilt 92, No. 2-3, 2003.
  24. Petkovic, M., ve Jonker, W., "Content based retrieval of spatio-temporal video events", **Proceedings of the IEEE International Workshop on Detection and Recognition of Events in Video**, 75-82, 2001.
  25. Petkovic, M., ve Jonker, W., "Cobra: a content based video retrieval system", **Proceedings of 8th Conference on Extending Database Technology - EDBT 2002**, LNCS, 2287, Springer-Verlag, 736-738, 2002.
  26. Petkovic, M., ve Jonker, W., **Content-based Video Retrieval: A Database Perspective**, Kluwer Academic Publishers, Boston, Monograph, ISBN 1-4020-7617-7 A.B.D., 2003.
  27. Roach, M., Mason, J., Xu, L., ve Stentford, F., "Recent Trends In Video Analysis: A Taxonomy Of Video Classification Problems", **Proceedings of the International Conference on Internet and Multimedia Systems and Applications**, IASTED, 348-354, 2002.
  28. Fischer, S., Lienhart, R., ve Effelsberg, W., "Automatic Recognition of Film Genres", **The 3rd ACM Int. Multimedia Conference and Exhibition**, 1995.
  29. Liu, Z., Huang, J., Wang, Y., ve Chen, T., "Audio feature extraction and analysis for scene classification", In **IEEE Signal Processing Society Workshop on Multimedia Signal Processing**, 1997.
  30. Liu, Z., Huang, J., ve Wang, Y., "Classification of TV programs based on audio information using hidden Markov Model", **IEEE Signal Processing Society Workshop on Multimedia Signal Processing**, 1998.
  31. Chang, C-W, ve Lee, S-Y, "A video information system for sport motion analysis", **Journal of Visual Languages and Computing**, Cilt. 8, 265-287, 1998.
  32. Colombo, C., Bimbo, A. D., ve Pala, P., "Retrieval of commercials by video semantics", **Computer Vision and Pattern Recognition - CVPR**, 572-577, 1998.
  33. Brand, J.D., Mason, J.S.D., ve Pawlweski, M., "Face detection in colour images", **International Conference on Image Processing**, 2001.
  34. Wang, H., ve Chang, S-F., "A highly efficient system for automatic face region detection in MPEG video", **IEEE Transactions on Circuits and Systems for Video Technology**, Cilt 7, No. 4, 1997.
  35. Hu, W., Tan, T., Wang, L. ve Maybank, S., "A survey on visual surveillance of object motion and behaviors", **IEEE Transactions on Systems, Man, and Cybernetics, Part C, Applications and Reviews**, Cilt 34, No. 3, 334-352, 2004.
  36. Aggarwal, J.K., ve Cai, Q., "Human motion analysis: A review", **Computer Vision and Image Understanding**, Cilt 73, No. 3, 428-440, 1999.
  37. Cohn, A. G., Magee, D. R., Galata, A., Hogg, D., ve Hazarika, S. M., "Towards an architecture for cognitive vision using qualitative spatio-temporal representations and abduction", **Proceedings of Spatial Cognition**, 232-248, 2003.
  38. Lavee, G., Rivlin, E., Rudzsky, M., "Understanding Video Events: A Survey of Methods for Automatic Interpretation of Semantic Occurrences in Video", **IEEE Transactions on Systems, Man, and Cybernetics, Part C, Applications and Reviews**, Cilt 39, No 5, 489-504, 2009.
  39. Micheloni, C., Snidaro, L., ve Foresti, G. L., "Exploiting temporal statistics for events analysis and understanding", **Image and Vision Computing**, Cilt 27, No 10, Special Section: Computer Vision Methods for Ambient Intelligence, 1459-1469, 2009.
  40. Oliver, N., Rosario, B., ve Pentland, A., "A Bayesian computer vision system for modeling human interactions", **IEEE Transactions Pattern Analysis Machine Intelligence**, Cilt 22, No 8, 831-843, 2000.
  41. Smeaton, A.F. ve Over, P., "TRECVID: Benchmarking the Effectiveness of Information Retrieval Tasks on Digital Video", **Proceedings of the 2nd International Conference on Image and Video Retrieval**, Urbana, Editör: E.M. Bakker, T.S. Huang, M.S. Lew, N. Sebe, ve X. Zhou, Springer-Verlag, London, 10-27, 451-456, 2003.
  42. Amir A., Argillander J., Campbell M., Haubold A., Iyengar G., Ebadollahi S., Kang F., Naphade M.R., Natsev A., Smith J.R., Tešić J., ve Volkmer T., "IBM Research TRECVID-2005

- Video Retrieval System”, **Proceedings of the TRECVID 2005 Workshop**, NIST Special Publications, Gaithersburg, MD, A.B.D., 2005.
43. Naphade, M.R., ve Smith, J.R., “On the Detection of Semantic Concepts at TRECVID”, **Proceedings of the ACM International Conference on Multimedia -MM '04**, ACM Press, New York, NY, 660-667, 2004.
  44. TRECVID, 2009, Video Retrieval Evaluation (TRECVID), Çevrimiçi: “[http://www-nlpir.nist.gov/projects/trecvid/](http://www.nlpir.nist.gov/projects/trecvid/)”.
  45. S. Ayache ve G. Quenot, “TRECVID 2007 collaborative annotation using active learning”, **Proceedings of TRECVID**, 2007.
  46. Huijismans, D.P. ve Sebe, N., “How to Complete Performance Graphs in Content-Based Image Retrieval: Add Generality and Normalize Scope”, **IEEE Transactions on Pattern Analysis and Machine Intelligence**, Cilt 27, No. 2, 245-251, 2005.
  47. Pickering, M.J. ve Rüger, S. “Evaluation of key-frame based retrieval techniques for video”, **Computer Vision and Image Understanding**, Cilt 92, No. 2, 217-235, 2003.
  48. Silva, G.C. De, Yamasaki, T., ve Aizawa, K., “Evaluation of video summarization for a large number of cameras in ubiquitous home”, **Proceedings of the 13th ACM international conference on Multimedia**, ACM, 820-828, Singapore, 2005.
  49. Foote, J., “An Overview of Audio Information Retrieval”, **ACM Multimedia Systems**, Cilt 7, No. 1, 42-51, 1999.
  50. Downie, J.S., “Toward the scientific evaluation of music information retrieval systems”, **Proceedings of the International Conference on Music Information Retrieval**, Baltimore, A.B.D., 25-32, 2003.
  51. CAVIAR websitesi, Context Aware Vision using Image-based Active Recognition, Çevrimiçi: “<http://homepages.inf.ed.ac.uk/rbf/CAVIAR/>”, 2009.
  52. VACE websitesi, Video Analysis and Content Extraction, Çevrimiçi: “<http://www.icarda.org/InfoExploit/vace/index.html>”, 2009.
  53. CREDS websitesi, Call for Real-Time Event Detection Solutions (CREDS) for Enhanced Security and Safety in Public Transportation, Çevrimiçi: “<http://www.dsp.elet.polimi.it/avss2005/CREDS.pdf>”, 2009.
  54. CLEAR websitesi, Classification of Events, Activities and Relationships – Evaluation Campaign and Workshop, Çevrimiçi: “<http://www.clear-evaluation.org/>”, 2009.
  55. PETS websitesi, IEEE International Workshop on Performance Evaluation of Tracking and Surveillance (PETS), Çevrimiçi: “<http://www.pets2006.net/>”, 2009.
  56. ETISEO: Video understanding Evaluation, Çevrimiçi: “<http://www.silogic.fr/etiseo/>”, 2009.
  57. SanMiguel, J. C., Martínez, J. M. ve Garcia, A., “An Ontology for Event Detection and its Application in Surveillance Video”, **Advanced Video and Signal Based Surveillance, Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance**, 220-225, 2009.
  58. i-LIDS websitesi, Image library for intelligent detection systems. Çevrimiçi: “<http://scienceandresearch.homeoffice.gov.uk/hosdb2/physical-security/detection-systems/i-lids/>”, 2009.
  59. Chen, M., Chen, S-C, ve Shyu, M-L, “Hierarchical Temporal Association Mining for Video Event Detection in Video Databases”, **22nd International Conference on Data Engineering Workshops**, 137-145, 2007.
  60. Bertini, M., Del Bimbo, A., ve Torniai, C., “Soccer Video Annotation Using Ontologies Extended with Visual Prototypes”, **International Workshop on Content-Based Multimedia Indexing - CBMI '07**, 212-218, 2007.
  61. Poli, R., Kameas, A., ve Seremeti, L., “Ontology and Multimedia”, **Encyclopedia of Multimedia Technology and Networking E-book**, Editör: Margherita Pagani. Idea Group Publishing, 2005.
  62. Harit, G., Chaudhury, S., ve Paranjpe, J., “Ontology guided access to document images”, **Proceedings Eighth International Conference on Document Analysis and Recognition**, Cilt 1, 292-296, 2005.
  63. Simou, N., Saathoff, C., Dasiopoulou, S., Spyrou, E., Voisine, N., Tzouvaras, V., Kompatsiaris, I., Avrithis, Y., ve Staab, S., “An Ontology Infrastructure for Multimedia Reasoning”, **Visual Content Processing and Representation, 9th International Workshop - VLBV 2005**, LNCS 3893, Editör: L. Atzori vd., Sardinia, İtalya, 51-60, 2006.
  64. Mallik, A., Pasumarthi, P., ve Chaudhury, S., “Multimedia ontology learning for automatic annotation and video browsing”, **Proceeding of the 1st ACM international conference on Multimedia information retrieval**, Multimedia browsing and summarization, Vancouver, British Columbia, Kanada, 387-394, 2008.
  65. W3C, The World Web Consortium (W3C), Çevrimiçi: “<http://www.w3.org/>”, 2009.
  66. Albanese, M., Maresca, P., Picariello, A., ve Rinaldi, A. M., “Towards a Multimedia Ontology System: an Approach Using TAO\_XML”, **In Proceedings of the Eleventh International Conference on Distributed Multimedia Systems - DMS 2005**, Banff, Canada, 52-57, 2005.
  67. Nevatia, R., Zhao, T., ve Hongeng, S., “Hierarchical Language-based Representation of Events in Video Streams”, **Conference on**

- Computer Vision and Pattern Recognition Workshop - CVPRW '03**, Cilt 4, 39-39, 2003.
68. Nevatia, R., Hobbs, J., ve Bolles, B., "An Ontology for Video Event Representation", **Conference on Computer Vision and Pattern Recognition Workshop - CVPRW '04**, 119-119, 2004.
  69. Francois, A.R.J., Nevatia, R., Hobbs, J., Bolles, R.C., ve Smith, J.R., "VERL: an ontology framework for representing and annotating video events", **IEEE Multimedia**, Cilt 12, No 4, 76 – 86, 2005.
  70. Voisine, N., Dasiopoulou, S., Mezaris, V., Spyrou, E., Athanasiadis, T., Kompatsiaris, I., Avrithis, Y., ve Strintzis, M.G., "Knowledge-assisted video analysis using a genetic algorithm", **Proceedings of 6th International Workshop on Image Analysis for Multimedia Interactive Services - WIAMIS 2005**, 2005.
  71. Vakali, A., Hacid, M-S., ve Elmagarmid, A., "MPEG-7 based description schemes for multi-level video content classification", **Image and Vision Computing**, Cilt 22, No 5, 367-378, 2004.
  72. Snoek, C., Worring, M., Koelma, D., ve Smeulders, A., "A Learned Lexicon-Driven Paradigm for Interactive Video Retrieval", **IEEE Transactions on Multimedia**, Cilt 9, No 2, 280-292, 2007.
  73. Bagdanov, A., Bertini, M., Bimbo, A. D., Serra, G., ve Torniai, C., "Semantic annotation and retrieval of video events using multimedia ontologies", **Proceedings of IEEE International Conference on Semantic Computing - ICSC**, Irvine, California, A.B.D., 713-720, 2007.
  74. Tannenbaum, A., **Metadata Solutions: Using Metamodels, Repositories, XML, and Enterprise Portals to Generate Information on Demand**, Boston, A.B.D., Addison-Wesley, 2002.
  75. Farrar, S., Lewis, W., ve Langendoen, T., "A Common Ontology for Linguistic Concepts", **Proceedings of the Knowledge Technologies Conference**, Seattle, Washington, A.B.D., 2002.
  76. Bouet, M., ve Aufaure, M-A., "New Image Retrieval Principle: Image Mining and Visual Ontology", **Multimedia Data Mining and Knowledge Discovery**, Bölüm 9, Editör: Valery A. Petrushin, Latifur Khan, 2007.
  77. Hollink, L., Worring, M. ve Schreiber, G., "Building a Visual Ontology for Video Retrieval", **ACM International Conference on Multimedia - MM'05**, 479-482, 2005.
  78. Bimbo, A. D., ve Bertini, M., "Multimedia Ontology Based Computational Framework for Video Annotation and Retrieval", **LNCS - Multimedia Content Analysis and Mining**, 18-23, 2007.
  79. Petridis K., Precioso, F., Athanasiadis, T., Avrithis, Y., ve Kompatsiaris, Y., "Combined Domain Specific and Multimedia Ontologies for Image Understanding", **Workshop on Mixed-reality as a Challenge to Image Understanding and Artificial Intelligence at the 28th German Conference on Artificial Intelligence - KI 2005**, Koblenz, Almanya, 2005.
  80. Dasiopoulou, S., Tzouvaras, V., Kompatsiaris, I., ve Strintzis, M.G., "Enquiring MPEG-7 based Multimedia Ontologies", **Multimedia Tools and Applications**, DOI: 10.1007/s11042-009-0387-4, 2009.
  81. Petridis, K., Anastasopoulos, D., Saathoff, C., Timmermann, N., Kompatsiaris, Y., ve Stabb, S., "M-OntoMat-Annotizer: Image Annotation Linking Ontologies and Multimedia Low-Level Features", **KES 2006, Part III**, LNAI 4253, Editör: B. Gabrys, R., J., Howlett, and L.C. Jain, 633-640, 2006.
  82. Naphade, M., Smith, J. R., Tesic, J., Chang, S-F, Hsu, W., Kennedy, L., Hauptmann, A., ve Curtis, J., "Large-Scale Concept Ontology for Multimedia", **IEEE MultiMedia**, Cilt 13, No. 3, 86-91, 2006.
  83. AIMatShape projesi websitesi, Çevrimiçi: "<http://www.aimatshape.net/>", 2009.
  84. Tsinarakis, C., Polydoros, P., ve Christodoulakis, S., "Interoperability support between MPEG-7/21 and OWL in DS-MIRF", **Transactions on Knowledge and Data Engineering - TKDE, Special Issue on the Semantic Web Era**, Cilt. 19, No 2, 219–232, 2007.
  85. Aldershoff, F., Salden, A. H., Iacob, S., ve Kempen, M., "Supervised multimedia categorization", **Proceedings of SPIE 5021**, 100, DOI:10.1117/12.476242, 2003.
  86. Kienast, G., Zeiner, H., Hofmair, P., Schlatte, R., Thallinger, G., Burger, T. vd., "Representation Techniques for Multimedia Objects", **SALERO**, Deliverable 3.1.1 IST FP6-027122, 2006.
  87. Bachler, M. S., Buckingham Shum, S. J., De Roure, D. C., Michaelides, D. T. ve Page, K. R., "Ontological Mediation of Meeting Structure: Argumentation, Annotation, and Navigation", **1st International Workshop on Hypermedia and the Semantic Web - HTSW2003**, Nottingham, İngiltere, 2003.
  88. Eleftherohorinou, H., Zervaki, V., Gounaris, A., Papastathis, V., Kompatsiaris, Y., Hobson, O., "Towards a Common Multimedia Ontology Framework (Analysis of the Contributions to Call for a Common Multimedia Ontology Framework Requirements)", Report on ACEMedia Project, version: 1.0, Date: 25-04-2006, 2006.
  89. SmartWeb projesi websitesi, Çevrimiçi: "[http://www.smartweb-project.org/ontology\\_en.html](http://www.smartweb-project.org/ontology_en.html)", 2009.
  90. Boll, S., ve Klas, W., "ZyX : A Multimedia Document Model for Reuse and Adaptation of Multimedia Content", **IEEE Transactions on**

- Knowledge and Data Engineering**, Cilt 13, No 4, 361 – 382, 2001.
91. Doulaverakis, C., Kompatsiaris, Y., ve Strintzis, M.G., "Ontology-Based Access to Multimedia Cultural Heritage Collections - The REACH Project", **The International Conference on Computer as a Tool, 2005 - EUROCON 2005**, Cilt 1, 151-154, 2005.
  92. Petridis, K., Bloehdorn, S., Saathoff, C., Simou, N., Dasiopoulou, S., Tzouvaras, V., Handschuh, S., Avrithis, Y., Kompatsiaris, Y., ve Staab, S., "Knowledge representation and semantic annotation of multimedia content", **IEEE Proceedings on Vision, Image and Signal Processing**, Cilt 153, No 4, 255-262, 2006.
  93. Petridis, K., Kompatsiaris, I., Strintzis, M. G., Bloehdorn, S., Handschuh, S., Staab, S., Simou, N., Tzouvaras, V., ve Avrithis, Y. S., "Knowledge Representation for Semantic Multimedia Content Analysis and Reasoning", **Proceedings of the European Workshop for the Integration of Knowledge, Semantics and Digital Media Technology - EWIMT 2004**, Editör: Paola Hobson ve Ebroul Izquierdo ve Ioannis Kompatsiaris ve Noel E. O' Connor, Londra, İngiltere, 2004.
  94. Little, S., Martinelli, M., Salvetti, O., Güdükbay, U., Ulusoy, Ö., Chalendar, G., ve Grefenstette, G., "Integration of Structural and Semantic Models for Multimedia Metadata Management", **Content-Based Multimedia Indexing – CBMI'07**, Bordeaux, Fransa, 40-45, 2007.
  95. Troncy, R., Celma, O., Little, S., Garcia, R., ve Tsinaraki, "MPEG-7 based Multimedia Ontologies: Interoperability Support or Interoperability Issue?", **Proceedings of 1st Workshop on Multimedia Annotation and Retrieval enabled by Shared Ontologies**, Genova, İtalya, 2-14, 2007.
  96. Garcia, R. ve Celma, O., "Semantic Integration and Retrieval of Multimedia Metadata", **5th International Workshop on Knowledge Markup and Semantic Annotation - SemAnnot'05**, Galway, İrlanda, 2005.
  97. Arndt, R., Troncy, R., Staab, S., Hardman, L., ve Vacura, M., "COMM: Designing a Well-Founded Multimedia Ontology for the Web", **6th International Semantic Web Conference - ISWC**, 2007.
  98. Tsinaraki, C., Polydoros, P., ve Christodoulakis, S., "Interoperability support for Ontology-based Video Retrieval Applications", **Proceedings of 3rd International Conference on Image and Video Retrieval - CIVR'04**, 582-591, 2004.
  99. Bertini, M., Bimbo, A. D., ve Serra, G., "Video Event Annotation using Ontologies with Temporal Reasoning", **Proceedings of IEEE Italian Research Conference on Digital Library Systems - IRCDL**, Padova, İtalya, 24-25, 2008.
  100. Lee, J., Abualkibash, M., H., ve Ramalingam, P., K., "Ontology-based Shot Indexing for video Surveillance System", **Innovations and Advanced Techniques in Systems, Computing Sciences and Software Engineering**, Editör: K. Elleithy, 237-242, 2008.
  101. Mezaris, V., Strintzis, M. G., "Object Segmentation and Ontologies for MPEG-2 Video Indexing and Retrieval", **LNCS 3115 - CIVR 2004**, Editör: P. Enser vd., 573-581, 2004.
  102. Jiang, Y.-G., ve Ngo, C.-W., "Bag-of-visual-words expansion using visual relatedness for video indexing", **ACM SIGIR'08, Conference on Research and Development on Information Retrieval**, Singapore, 769-770, 2008.
  103. Benmokhtar, R., ve Huet, B., "Hierarchical Ontology-Based Robust Video Shots Indexation Using Global MPEG-7 Visual Descriptors", **Seventh International Workshop on Content-Based Multimedia Indexing**, 195-200, 2009.
  104. Bertini, M., Bimbo, A. D., ve Serra, G., "Learning ontology rules for semantic video annotation", **Proceedings of ACM International Conference on Multimedia, Many Faces of Multimedia Semantics - MS**, Vancouver, British Columbia, Kanada, 1-8, 2008.
  105. Ballan, L., Bertini, M., Bimbo, A. D., ve Serra, G., "Semantic Annotation of Soccer Videos by Visual Instance Clustering and Spatial/Temporal Reasoning in Ontologies", **Multimedia Tools and Applications**, DOI: 10.1007/s11042-009-0342-4, 2009.
  106. Song, D., Liu, H. T., Miyoung, C., Hanil, K., ve Pankoo, K., "Domain knowledge ontology building for semantic video event description", **CIVR 2005: international conference on image and video retrieval**, Editör: W.-K. Leow et al., Cilt 3568, 267-275, No. 4, Singapore, 2005.
  107. Hunter, J., "Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology", **Proceedings of 1st International Semantic Web Working Symposium - ISWC**, 261–281, 2001.
  108. OAI: Open Archives Initiative websitesi, Çevrimiçi: "<http://www.openarchives.org/>", 2009.
  109. TV-Anytime Forum, Çevrimiçi: <http://www.tv-anytime.org/>, 2009.
  110. Tsinaraki C., Polydoros P., ve Christodoulakis S., "Integration of OWL ontologies in MPEG-7 and TV-Anytime compliant semantic indexing", **Proceedings of the 16th International Conference on Advanced Information Systems Engineering - CAISE**, Riga, Latvia, 398-413, 2004.
  111. NewsML websitesi, Çevrimiçi: "<http://www.newsml.org/>", 2009.
  112. Yaginuma, T., Pereira, T. S. M., ve Baptista, A. A., "Metadata Elements for Digital News Resource Description", **Proceedings of the III**

- Congresso Luso-Moçambicano de Engenharia**, Maputo, Mozambik, 2003.
- 113.**O'Connor, N., Cooke, E., le Borgne, H., Blighe, M., ve Adamek, T., "The ace-toolbox: low-level audiovisual feature extraction for retrieval and classification", **2nd IEEE European Workshop on the Integration**, 2005.
- 114.**The Protégé projesi websitesi, Çevrimiçi: "http://protege.stanford.edu", 2009.
- 115.**IBM MPEG-7 Etiketleme aracı websitesi, Çevrimiçi:"http://www.alphaworks.ibm.com/tech/videoannex",2009.

