



طراحی و پیاده سازی یک زبان پرسش روی بانک های ویدئویی با تاکید بر فاکتورهای زمانی

پریناز خوشبخت نژاد، فوق لیسانس کامپیوتر گرایش نرم افزار مدرس حق التدریس دانشگاه آزاد اسلامی، واحد خوی
pkhoshbakhtnezhad@yahoo.com
دکتر محمدرضا کنگاوری، kangavari@iust.ac.ir

1. چکیده

پیشرفت های حاصل در زمینه فشرده سازی مالتی مدیا، استانداردهایی را به همراه افزایش کارایی شبکه های کامپیوتری و رسانه های ذخیره سازی باعث پشتیبانی بخش وسیعی از مالتی مدیا شده است. اما، ایجاد بانک های اطلاعاتی بزرگ جهت دسترسی آن به منابع، فرایند جستجو و بازیابی اطلاعات را پیچیده کرده است. که دلیل اصلی این مشکلات مخفی بودن معنای موجود در اینگونه داده ها و گستردگی دامنه معنایی آنها می باشد، خصوصاً عامل زمان نقش عمده در این گستردگی دامنه معنایی آنها به عهده دارد. زبانهای حاضر بدون در نظر گرفتن عامل زمان به عنوان یک عامل کلیدی در درک معنایی تصاویر ویدئویی، طراحی و کدام از نقاط ضعف دیگری رنج می برند.

برای غلبه بر این مش دستیابی به زبانی که بتواند عامل زمان را در اختیار کاربر جهت طرح پرسش و دریافت اطلاعات بگذارد، مطرح شده است. زبان پرسشی حاضر، در تکمیل یک معماری جدید، برای پایگاه داده ی ویدئویی، بکار رفته است

کلمات کلیدی:

هستان شنا

روی پایگاه داده های ویدئویی

2. مقدمه

پایین تصویر و میزان شباهت است که در همگی آنها به خوبی کمبود ادراک و معنی دیده می شود.

در خیلی مواقع پرس و جو مبتنی بر مثال با پرس وجوی مبتنی بر حاشیه نویسی را با هم ترکیب میکنند. برای مثال سیستم [15] [2] [13] به کاربر اجازه پرس وجوهای ترکیبی را می دهد. [17] در پرس وجوهای ترکیبی از ویژگیهای هر دو را با رنگ برجسته بکار می برد. سیستمهای دیگر که قادرند از این روش استفاده کنند عبارتند از MOOL که میتواند از آن برای بازیابی محتوای اسناد وبازیابی محتوای رسانه استفاده کرد.

امروزه جستجوگر مبتنی بر محتوا شیء های مالتی مدیا بیشتر مورد توجه قرار گرفته تا

در دو دهه اخیر پیشرفت های قابل توجهی در مالتی مدیا و تکنولوژی ذخیره سازی که منجر به ایجاد انبار بزرگی از تصاویر، ویدئو و داده صوتی دیجیتال شده است. اما جستجو برای مالتی مدیا آسان نیست چراکه داده مالتی مدیا مخالف متن است و نیازمند مراحل زیادی از پیش پردازش است تا همچنان تصمیم گیری مناسبی برای پردازش کردن داشته باشد

با وجود تحقیقات فراوان در این زمینه، خروجی چنین تحقیقاتی تنها محدود به تعدادی سیستم یا پرس و جوهای مبتنی بر مثال¹ [20]، [21]، [23] بوده است که عمدتاً مبتنی بر ویژگیهای سطح

ویژگی‌های ویژوالی چون رنگ، الگو، شکل و اطلاعات فضایی² که در وجود مناظر ویژوال هستند. این ویژگی‌ها بکار گرفته می‌شوند تا تشابه بین تصاویر مختلف و ویدیوها را محاسبه کند [8].

مجموعه زبانهای پرس وجو معنانشناسی³ چون [12]، [9] و OQL جهت توسعه زبانهای پرس وجو پایگاه کلاسیکی SQL پیشرفته مطرح شده اند. مشکل عمده ای که در این زبانهای پرس وجو دیده می‌شود این است هیچکدام از این ها بر پایه نیازمندی‌های کاربران بررسی نشده است. هر کدام از این زبانها بتنهایی یک زیرمجموعه از ساختار توانمند و توصیفات مبتنی بر محتوا فوکس شده که بر پایه این است که کاربران می‌توانند محتوای سعی بصری را پرس وجو کنند.

مدل‌های داده ویدئویی که روی محتوای معنا شناسی⁴ پایه‌ریزی شده اند [3]، [4]، [6]، [10]، [11]، [14]، [16]، [18]، [19] به خوبی می‌توانند از یک پرس وجو روشن پیش‌تیبانی کنند. اما باید قسمتی از آن روی حاشیه‌نویسی دستی تکیه کند و محدودیت دیگر محتوای معنا شناسی در مبهم بودن و وابسته به زمینه بودن است. این محدودیت می‌تواند با محدود کردن زمینه⁵ تولید چنین توصیف معنا شناسی برای انواع کاربرد⁶ها کنترل شود و مشکل دیگر سیستم‌هایی که بر پایه بازیابی معنانشناسی⁷ها استوار شده شکاف معنایی است.

در راستای تحقیقات انجام شده پروژه MUMIS⁸ نیز از خروجی اطلاعات مبتنی بر هستان شناسی استفاده می‌کند تا نتایج

بازیابی اطلاعات را در آرشیوهای مالی مدیا بهبود دهد [22].

این پروژه مسائل بازیابی اطلاعات معنانشناسی در بعد مالی مدیا را آدرس دهی میکند. تکنیک های آدرس آن عبارتند از استخراج اطلاعات، شناسایی گفتار اتوماتیک و استخراج فریم کلیدی از محتوای ویدئو تا درستی چند زبانه را برای آرشیوهای مالی مدیا آسان کند. به علاوه پروژه MUMIS شامل یک ماژول است حاشیه نویسی های استخراج شده از منابع مختلف را در یک اجتماع خیلی کامل، توصیف اولیه از محتوایشان را ترکیب کند.

MUMIS نیز در یک هستان شناسی با یک فرمت XML خیلی ساده کد گذاری شده است و بطور کامل می‌تواند با فرمت RDF یا OWL نیز بیان شود، و این پروژه می‌تواند از سه زبان مختلف پشتیبانی کند.

فرق MUMIS با پژوهش ما:

1. پروژه MUMIS بیشتر روی استخراج اطلاعات و حاشیه نویسی آن بررسی شده و نسبتاً اهمیت کمتری به زبان پرس وجو داده است. درحالیکه این پژوهش با در نظر گرفتن حاشیه نویسی و استفاده از واژه های زمانی در هستان شناسی وابسته به زمان می‌تواند پرس وجوی قابل قبولی را برای کاربر ایجاد کند.

2. ساختار و اسلوب آن برای بازیابی نویسی ها در اسلوب XML با پژوهش ما فرق دارد.

3. MUMIS نیز در یک هستان شناسی با یک فرمت XML خیلی ساده کد گذاری شده است و بطور کامل می‌تواند با فرمت RDF یا OWL نیز بیان شود در حالیکه این سیستم هیچ گونه پشتیبانی از موارض فوق ندارد.

4. در این پژوهش ما برای استخراج اطلاعات و پرس وجو از هستان شناسی وابسته به زمان استفاده کردیم در حالیکه

² Spatial

³ Semantic

⁴ Semantic Content

⁵ Context

⁶ Application

⁷ Semantic

⁸ Multi-Media Indexing and Searching

5. نگاه کلی بر سیستم پیشنهادی

شکل بعدی معماری کلی چارچوب پیشنهادی را نمایش می دهد. هدف در اینجا طراحی يك زبان پرسشی سطح بالا است که عامل زمان امکان دستیابی به يك درک بالا و مناسب از داده های ویدئویی را فراهم می آورد. نمایه وابسته به زمان برای داده های ویدئویی، در قالب هستان شناسی وابسته به زمان است. اما جهت لمس خروجی سیستم و مثالی از کاربرد آن و بعضی از پیمانه ها بدین منظور اضافه شده اند.

هسته چارچوب پیشنهادی را پیمانه پرس وجو، پیمانه ساخت هستان شناسی وابسته به زمان و پیمانه واسط کاربری، تشکیل می دهد. ولی در کل کارهای انجام شده در [1]، [5] ایسن سیستم، ورودی را به صورت داده خام ویدئویی وارد میکند. کاربر سطح پایین (مدیر سیستم) می بایستی دانش زمینه ای را به سیستم معرفی نماید. این کار به کمک تعریف هستان شناسی های مناسب و استفاده از هستان شناسی های موجود و وارد کردن آنها در سیستم صورت گیرد. بدین منظور يك واسط کاربر گرافیکی طراحی شده است. کاربر در يك ساختار ساده مفاهیم، وقایع و اشیاء مورد توجه خود را به سیستم معرفی می نماید. دانش فراهم آمده به شکل مجموعه ترم های مجاز در اختیار پیمانه پیش پردازنده قرار می گیرد. به عبارت دیگر مجموعه ترمهای مجاز برای حاشیه نویسی از این طریق به سیستم معرفی می شود.

پیمانه پیش پردازنده به کمک دانش زمینه ای معرفی شده و ابزارهای قالبی داده های ویدئویی، و فرآیندهایی که توسط کاربر به سیستم معرفی می شود، حاشیه نویسی خودکار را بر روی داده های ورودی انجام می دهد. پیمانه ساخت هستان شناسی وابسته به زمان، از حاشیه نویسی های فراهم شده به همراه اطلاعات زمانی، هستان شناسی وابسته به زمان متناظر با این حاشیه نویسی ها را برای يك شی داده ویدئویی ورودی می

پروژه MUMIS نامی از هستان شناسی وابسته به زمان به میان نیاورده است.

در پژوهش پیش رو قصد بر این است تا با استفاده از زبان پرس وجو⁹، چارچوب کلی برای ارائه بازیابی محتوای وابسته به زمان برای داده های خام ویدئویی ارائه شود. و به طور مختصر این پژوهش نسبت به زبانهای پرس وجوی قبل در موارد زیر جدید تر است:

1. دستیابی به زبان پرسشی جدید و سطح بالا روی بانکهای داده ویدئویی.
2. در نظر گرفتن عامل زمان به عنوان يك عامل کلیدی در درک معنای واقعی روی بانکهای داده ویدئویی.

3. هدف پروژه

هدف در اینجا طراحی يك زبان پرسشی سطح بالا است که عامل زمان، امکان دستیابی به يك درک بالا و مناسب از داده های ویدئویی را فراهم می آورد و در طی آن ما زبان پرسش خود را ¹⁰TOQL گذاشتیم. نمایه وابسته به زمان برای داده های ویدئویی، در قالب هستان شناسی وابسته به زمان است و با توجه به این نگرش عامل زمان می تواند در اختیار کاربر، جهت طرح پرسش و دریافت اطلاعات در نظر گرفته شود.

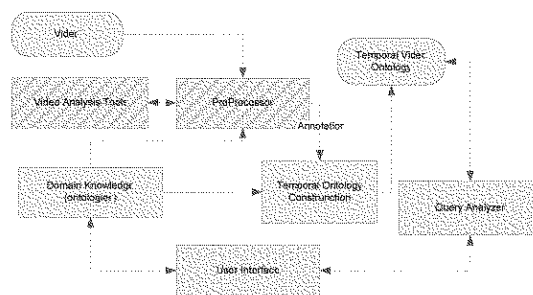
4. دستاوردها

- 1- حل برخی از مشکلات موجود در سیستم های مشابه
- 1-1 در نظر گرفتن داده های ویدئویی به عنوان مستند وابسته به زمان
- 2-1 حل مشکل تعداد کلمات محدود در پرس وجو
- 2- انعطاف پذیری و قابل انعطاف بودن آن
- 3- گستردگی دامنه معنایی آن برای پرس وجو

⁹ Query Language

¹⁰ Temporal Ontology Query Language

سازد. به عبارت دیگر فرا داده های فراهم آمده برای داده های ویدئویی ورودی، به شکل یک هستان شناسی وابسته به زمان نمود پیدا می کند.



شکل 1. بلاک دیاگرام سیستم پیشنهادی و رابطه میان پیمانه های مختلف

حال در این پژوهش با فرض وجود این پیمانه ها کاربر می تواند به کمک یک واسطه گرافیکی، پرس و جوهای وابسته به زمان از مفاهیم تعریف شده در دانش زمینه ای را بازیابی و مشاهده نماید. توجه به این نکته ضروری است که سیستم پیشنهادی چارچوبی را برای هر یک از برنامه های کاربردی مورد علاقه فراهم می آورد که در طی بازیابی، امکان کار با مفاهیم سطح بالا و معنایی وابسته به زمان وجود دارد. علاوه بر این ارائه مفاهیم زمانی، امکان استنتاج وابسته به زمان بر روی مفاهیم را فراهم می آورد. در ادامه هر یک از پیمانه های سیستم تشریح شده است.

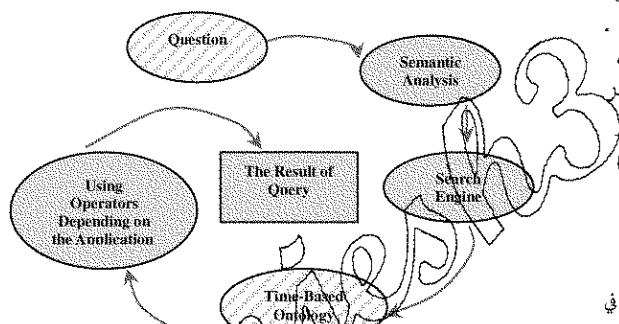
6. توصیف زبان پرسشی

در این پژوهش از ساختار معرفی شده در فصل سوم یعنی هستان شناسی وابسته به زمان، برای ارائه یک زبان پرس و جو وابسته به زمان استفاده شده است. همانطور که اشاره شد، سیستم از پیمانه های متعددی تشکیل شده است که هر یک وظیفه خاصی را بر عهده دارند. زبان پرسشی ما این امکان را به کاربر می دهد تا علاوه بر پرس و جو روی اشیاء، مفاهیم و کلیه وقایع مورد قبول در یک پایگاه داده ویدئویی، عامل زمان را نیز به عنوان یک اصل کلی در نظر بگیرد.

شکل 2 مراحل توصیف زبان پرسشی سیستم را نمایش می دهد. توجه به این نکته ضروری است که هستان شناسی وابسته به زمان و طرح جمله پرسشی کاربر به عنوان ورودی برای سیستم ما تلقی می شود. بحث مرتبط با هستان شناسی وابسته به زمان در [1] به تفصیل شرح داده شده است بنابراین از توضیح آن در اینجا صرف نظر می شود.

6-1. تجزیه و تحلیل معنایی

جمله پرسشی، با توجه به نوع کاربرد و انتزاع آن از زمان توسط کاربر مطرح می شود. در اینجا کاربر می بایست مشخص کند که به کدام رویکرد علاقمند و چه سطحی از دانش دانگی در انتزاع از زمان را در نظر گرفته است. به عنوان مثال لحظه یک واقعه برای وی بیشتر اهمیت است یا ساعت آن یا هر دو. حال در این پژوهش سعی ما در ابتدا فهماندن لغات و درک معانی کاربر به سیستم بوده تا سیستم با توجه به تقاضای کاربر پاسخ گو باشد.



شکل 2. چگونگی گردش داده ها و مراحل انجام کار برای پاسخ گویی به یک زبان پرسش وابسته به زمان برای یک داده ویدئویی. مراحل کلی که در شکل به شکل هاشور خورده دیده میشوند به عنوان ورودی سیستم ما تلقی می شوند. در نهایت نتیجه پرس و جو به عنوان نتیجه پردازش ایجاد خواهد شد.

```

}
Store query as Ontology based
Time
}
}

```

8. تست سیستم بر روی داده واقعی

همانطور که پیشتر نیز گفته شد دانش زمینه ای تعریف شده برای ویدئوهای فوتبال شامل وقایع، اشیاء و مفاهیم است. وقایع شامل گل، خطا، کرنر، تعویض و شوت به سمت گل است. مفاهیم تعریف شده شامل برد، باخت، تساوی، نتیجه بازی است. اشیاء تعریف شده شامل تیم، داور، استادیوم محل برگزاری و بازیکن است. پس هر یک از عناصر ارائه شده در هستان شناسی دانش زمینه ای دارای خصیصه های خاص خود همانند خصیصه های زمانی است. حاصل انجام عملیات هستان شناسی وابسته به زمان در قالب فایل XML به عنوان ورودی برای سیستم ما داده شده است. قسمتی از هستان شناسی وابسته به زمان در فصل قبل نمایش داده شده است. حال سیستم ما با توجه به دو ورودی و پردازشگرهای موجود خود شروع به پاسخ گویی به سوالات کاربر میکند. سوالاتی چون؛ چند گل تا دقیقه 12 ام زده شده است؟ چه اتفاقاتی در دقیقه 11 ام رخ داده؟ چند بار خطا صورت گرفته؟ آیا تیم A به تیم B دو گل زده است؟ در چه زمان هایی تیم A گل زده است؟ و غیره. تمامی این سوالات بصورت اختیاری با هر بازه زمانی دخواه توسط کاربر، از سیستم می تواند پرسیده شود و هیچگونه محدودیتی در تعداد سوالات و نحوه پرسش وجود ندارد.

9. نتیجه گیری و سمت و سوی تحقیقات آینده

نتایج بدست آمده نشان دهنده توانایی سیستم در ارائه زبان های پرس وجوی وابسته به

2-6. موتورهای جستجو گر

حاصل کار پژوهش [1]، معرفی چارچوبی برای استخراج معانی و ارائه آن به شکل هستان شناسی وابسته به زمان است و مهمترین مزیت آن اجتماع اطلاعات زمانی با مفاهیم و معانی استخراج شده از ویدئو است. پس هستان شناسی وابسته به زمان به عنوان یک ساختار هستان شناسانه که دانش وابسته به زمان را در حوزه ای خاص و بر پایه دانش زمینه ای و هستان شناسی زمان شناسانده می شود. حال سیستم ما با توجه به این موضوع پس از تجزیه و تحلیل معنایی و درک کلیه معانی لغات، شروع به جستجو در ورودی دیگر ما یعنی هستان شناسی وابسته به زمان می کند و در صورت نیاز این موتور جستجو گر عناصر موجود در هستان شناسی وابسته به زمان را با هم ادغام می کند تا با استنتاج خود از عناصر به یک نتیجه مطلوب برسد.

7. الگوریتم بکار رفته برای زبان پرسشی مبتنی بر زمان

```

Prepare the query language
Use User Interface to get the query
language
If Qur(a[i] != 0){
    Do Statement Detection
    Do Text and Time Detection
    While ( Ontology based
    Time.hasLength ) {
        Get the Ontology based
    Time {
        // Use External Ontology based Time
        Processing Unit to Get Data
        Do Text Detection
        Do Time Detection
        Do Text Extraction
    }
    Make the query language {
        Read query language
        Read Ontology based
    Time
        Make Expressions to get
    data from Ontology based Time
        Get query language
    }
}

```

2. پرس وجو روی خود ویدئو
3. پرس و جو برای چندین بازی ورزشی مثل بسکت بال، هند بال و...
4. استفاده از منطق فازی جهت پرسش های نامعین¹¹ زمانی.

10. مراجع

- [1] بهرنگ قاسمی زاده، "ارائه چارچوبی برای استخراج هستان شناسی وابسته به زمان از داده های چند رسانه ای (ویدئو) حاشیه نویسی شده"، پایان نامه کارشناسی ارشد، دانشگاه علم و صنعت، دانشکده کامپیوتر، 85
- [2] M. Petkovic, W. Jonker. **CONTENT-BASED RETRIEVAL OF SPATIO-TEMPORAL VIDEO EVENTS.** *Computer Science Department, University of Twente P.O. Box 217, 7500 AE Enschede, The Netherlands.*
- [3] E.Bertino, F.Rabitti, and S.Gibbs. **Query processing in a multimedia document system.** *ACM Transactions on Office Information Systems.* 6(1):1-41, January 1988.
- [4] A.F.Cardenas, I. T. leong, R. K. Taira, R. Barker, and C.M.Breant. **The knowledgebased object-oriented PICQUERY+ language.** *IEEE Transactions on Knowledge and Data Engineering.* 5(4):644-657, August 1993.
- [5] راستگوی حقی، هادی، "روشی برای استخراج حاشیه نویسی خودکار از داده های ویدئویی"، پایان نامه کارشناسی ارشد، دانشگاه علم و صنعت، دانشکده کامپیوتر، 85
- [6] W. W. Chu, I. T. leong, and R. K. Taira. **A Semantic Modeling approach for image retrieval by content.** *The VLDB Journal.* 3:445-447, 1994.
- [7] N. Dimitrova and F.golshani. **EVA: A query language for multimedia information systems.** *In Proc Of Multimedia Information Systems, Tempe, Arizona, February 1992.*
- [8] M. Egenhofer. **Spatial SQL: A Query and Presentation Language.** *IEEE Transactions on Knowledge and Data Engineering.* 6(1): 86-95, January 1994.
- [9] E.Oomoto and K. Tanaka. **OVID: Design and implementation of a video-object database system.** *IEEE Transactions on Knowledge and Data Engineering.* August 1993.

زمان از داده های ویدئویی است. همانطور که پیشتر ذکر شد مهمترین ویژگی چارچوب پیشنهادی ارائه زبان پرس وجویی با اطلاعات زمانی مجتمع شده با معانی متضمن ویدئو در یک چارچوب هستان شناسانه وابسته به زمان است. این چارچوب، توانایی استفاده از هستان شناسی وابسته به زمان و امکان یک بازیابی هوشمندانه از داده های چند رسانه ای را میدهد. از نقاط قوت این زبان، در دستیابی به زبان پرس وجویی جدید سطح بالا و در نظر گرفتن عامل زمان به عنوان یک عامل کلیدی در درک معنایی روی بانکهای داده ویدئویی است. در حال حاضر هستان شناسی وابسته به زمان از زبانهای وب معنایی متداولی همانند RDF و OWL پشتیبانی نمی کند. در این پژوهش ما با استفاده از هستان شناسی وابسته به زمان، می توانیم امکاناتی جهت ارائه بازیابی دانش برای زمینه های مختلف فراهم کنیم. مزیت دیگر این زبان پرس وجو در امکان استفاده از هستان شناسی وابسته به زمان بر روی وب معنایی است. چرا که صدور نمایه بدست آمده به عنوان مدل معنایی وابسته به زمان از داده های ویدئویی به یکی از زبانهای وب معنایی سبب فراهم آوردن امکان استفاده از آنها را در وب معنایی آینده بوجود می آورد. می توان در پیاده سازی این سیستم، از روش های گوناگون هوش مصنوعی همانند منطق فازی جهت بهبود هر چه بیشتر نتایج بدست آمده، بهره گرفت. سیستم ارائه شده در صورت بسط مناسب می تواند به عنوان یک زبان پرس وجویی جهت پرس وجو از داده های ویدئویی مورد استفاده قرار گیرد. از آنجائیکه این سیستم برای توسعه دادن ساختار این زبان پرسش محدودیتی ندارد لذا برای کارهای آینده این پژوهش، موارد زیر را در نظر داریم:

1. چند زبانه کردن آن

[23] Chia-Hung Wei and Chang-Tsun Li. **Design of Content-based Multimedia Retrieval.** Department of Computer Science University of Warwick.

[10] S. Adali, K. S. Candan, S.-S. Chen, K. Erol, and V. S. Subrahmanian. **Advanced video information system: Data structures and query processing.** ACM-Springer Multimedia Systems Journal, 1996.

[11] T.-S. Chua and L.-Q. Ruan. **A video retrieval and sequencing system.** ACM. Transactions on Information Systems, 13(4), 1995.

[12] Q. Li, Y. Yang, W-K. Chung, "**CAROL: Towards a Declarative Video Data Retrieval Language**". Proc. of SPIE, Vol. 3561, 1998.

[13] J. Bach et. al. **The virage search engine: An open framework for image search engine.** In SPIE conference on Storage and retrieval of Image and video databases, 1997.

[14] C. Declerck and M.-S. Hacid. **A database approach for modeling and querying video data.** In IEEE Data Engineering, Australia, 1999.

[15] M. Petkovic, W. Jonker. **Content-based video retrieval by integrating spatio-temporal and stochastic recognition of events.** Computer science department, university of Twente, the Netherlands 2002.

[16] H. Jiang, D. Montesi, and K. Elmagarmid. **Videotext database systems.** In IEEE Int'l Conf. on Multimedia Computing and Systems, Ontario, Canada, June 1997

[17] E. Ardizzone, M. La Cascia, D. Molinelli, "**Motion and Color Based Video Indexing and Retrieval**", Int. Conf. on Pattern Recognition, CPR, Wien, Austria, Aug. 1996.

[18] R. Weiss, A. Duda, and D. Giord. **Content-based access to algebraic video.** In IEEE Int'l Conf. on Multimedia Computing and Systems, Boston, USA, 1994.

[19] Ankush Mittal. **An Overview of Multimedia Content-Based Retrieval Strategies.** Deptt. of Electronics and Computer Engg, Indian Institute of Technology, Roorkee, India, October 14, 2005.

[20] M. L. Cascia and E. Ardizzone. **JACOB : Just a content-based query system for video databases.** In IEEE Int. Conf. on Acoustics, Speech and Signal Processing, May 1996.

[21] Samuel Johnson. **QUERY-BY-EXAMPLE (QBE).**

[22] D.Reidsma¹, J. Kuper¹, T.Declerck², H.Saggion³, and H.Cunningham³. **Cross Document Ontology based Information Extraction for Multimedia Retrieval.** ¹ University of Twente, Dept. of Computer Science ² DFKI GmbH, Language Technology Lab. ³ University of Sheffield, Dept. of Computer Science.