Multimedia has been one of the fastest-growing, most dynamic fields of computer science. With the unrelenting flood of multimedia data and applications, the technical challenges keep coming. At the core are difficult and intriguing challenges across networking, information retrieval, content recognition, content protection, data management, and more. But while the field continues to take on new problems, what has happened with the old problems? To gain some insight, I counted the number of published articles on various multimedia topics during the 2001–2011 period by searching Google Scholar. Figure 1 shows the topics and search result counts.

The early days of multimedia laid out much of the scaffolding around which the field is still built. One of the first topics was video on demand, which in the late 1980s was envisaged as a futuristic broadband service enabled by high-bandwidth fiber optic cabling to homes. W. David Sincoskie described video on demand as a service “provided from a centralized location over a digital network ... similar to the currently popular videotape rental services.”¹ He divided the problem into two parts: communications and database access. Although he claimed that the communications challenge could be addressed with sufficient bandwidth, he predicted that the database access problem could take up to 10 years solve.

What followed was a flood of work on video compression, video data management, video servers, networking, and communication that have made video on demand routine today. Although research still continues today, beginning in 2008, there has been a year-over-year decrease in the number of video on demand technical articles, as Figure 2 shows. Since video on demand is one of the genuine successes in multimedia, this most likely indicates that many of the technical challenges have indeed been solved.

However, other notable multimedia topics have not been as successful. Face recognition is one such example. The basic face recognition problem was addressed as far back as the 1970s with Takeo Kanade’s thesis on computer processing for recognizing human faces, for which he processed 800 photographs and conducted experiments involving the identification of 20 people.² The next notable highlight did not come for another two decades with Matthew A. Turk and Alex P. Pentland’s appearance-based holistic approach based on eigenfaces.³

Although face recognition has been the subject of substantial work since then, there has been a marked decline in the number of new face recognition technical papers since 2008 (see Figure 2). A similar situation can be seen with another computer vision challenge: object

---


tracking and detection. Yet, no one can conclude that face recognition or object tracking and detection are solved problems today. Rather, the decreasing focus might reflect fatigue. Or, it could be that researchers have been regrouping for the next big two-decade breakthrough!

There are many other topics that produced bursts of activity. MPEG-7 resulted a notable period of research that roughly tracked the work of the MPEG standards body, as can be seen in Figure 2. When the MPEG-7 standard was finalized around 2004, the rate of technical articles began to slow. Although this decrease can be expected, it is intriguing that there are still many new MPEG-7 papers today. This could mean that the underlying problems are yet to be solved. Indeed, today we still have no standard metadata scheme for multimedia data.

Many long-standing multimedia topics are still gaining focus, as Figure 3 shows, including

---

**Figure 2. Year-to-year change in the number of new published articles on six multimedia topics.**

**Figure 3. Change in the number of published articles on various multimedia topics between 2010 and 2011.**
content-based retrieval, video surveillance, multimedia semantics, and event detection. Major breakthroughs are still needed to solve these problems, but the research community is still finding ways to make progress. Interestingly, shot detection is finding increasing focus. The research community thought that the underlying technical challenge—to automatically detect camera breaks in video—had been solved. However, newer versions of the problem have appeared in the context of today’s unconstrained online video that is not professionally edited.

A few, fast-growing multimedia topics don’t have a long track record. Figure 3 shows that social media and multimedia crowdsourcing are growing at a tremendous rate. Both reflect the recent interest in on-line social networks and are good examples of how multimedia is still finding a central role despite the many problems gone by.

**References**


**John R. Smith** is a senior manager of Intelligent Information Management at IBM T.J. Watson Research Center. Contact him at jsmith@us.ibm.com.

Selected CS articles and columns are also available for free at [http://ComputingNow.computer.org](http://ComputingNow.computer.org).