MARK EVERINGHAM was a brilliant colleague. You may have been aware of him at conferences where he asked penetrating questions that could crystallize a key aspect of a paper. In conversation with him you might have been stunned by a new connection he made between areas of research or to a crucial related work. These questions and observations were a reflection of his very broad knowledge and deep understanding of computer vision and machine learning. To us, they demonstrated his intellect and insight, but to him they were just a way of being helpful, a way of ensuring that the field made progress.

Mark was incredibly generous with his time. Those of us that worked with him are aware of how much he contributed behind the scenes, without expecting any recognition. Nowhere is this more apparent than in the organization of the PASCAL Visual Object Classes (VOC) challenge to which he devoted colossal amounts of time and effort. There are also the more visible contributions to the community in area chair duties at both CVPR and ECCV, as a program cochair for BMVC, and as a member of the TPAMI editorial board. Everything he did: research, experimentation, software, paper writing, talks, was of the highest standard and a testament to his intellectual stamina.

He was kind and demonstrated a gentle, dry wit that made time spent with Mark both stimulating and enjoyable.

Mark was born in Bristol in 1973, winning a scholarship to Clifton College, and completing his A levels at Filton College in 1991. Directly after school he worked on a research project for the Bristol Eye Hospital, developing software for remote electrodiagnosis. He continued his involvement with the project after heading to the University of Manchester to study computer science, winning prizes for top achievement every year, and was duly awarded the BSc with 1st class honors and the Williams-Kilburn medal for exceptional achievement in 1995. Returning to Bristol, he completed work on the electrodiagnosis project, leading to his first publication, in the journal *Electroencephalography and Clinical Neurophysiology* in 1996.

In 1997 he began his doctoral studies at the University of Bristol, supervised by Barry Thomas and Tom Troscianko, on mobile augmented reality aids for people with severe visual impairments. By presenting an enhanced image to the wearer’s visual system, users with low vision would be freed from the need for external assistance in many tasks. The approach Mark took was, as usual, based on a deep rethinking of the problem. Rather than attempting to enhance the image by emphasizing edges, he proposed to identify the semantic content of an image such that images could be enhanced in a content-driven manner, and to enhance regions rather than edges. This work led him to look at region segmentation algorithms, and there he discovered the difficulties of evaluating computer vision algorithms, leading to some of the most significant papers from his PhD. In particular, “Evaluating Image Segmentation Algorithms Using the Pareto Front,” presented at ECCV 2002, showed the importance of the choice of evaluation metric in a compelling way, and is notable for its inclusion of the “embarrassingly simple” baseline method of dividing the image into blocks, which sometimes appears on the Pareto front. In presentations of this work, Mark would draw out the humor in this fact, but also use it as a point of reference to illustrate the behavior of metrics, to give insight into the criteria, and ultimately to convince you that you had learned something.

Graduating with his PhD in 2002, Mark moved to Andrew Zisserman’s group at Oxford University’s Department of Engineering Science, where he worked on three projects which explored the level of supervision required for visual classification and detection tasks. The first aimed to detect and identify actors in relatively low resolution video footage, such as TV material from the 1970s. The method involved quite strong supervision where a 3D head and face model were built for each character (from images). These 3D models were then used to render images to train a discriminative tree-structured classifier which was then used as a sliding window detector. This person-specific

approach succeeded in detecting characters over a wide range of poses (frontal/profile) and scales, as described in the ICCV 2005 paper “Identifying Individuals in Video by Combining Generative and Discriminative Head Models.”

The second project had a similar aim of automatically identifying characters in video footage, but now with reduced human supervision. Here, the supervision was provided by using the subtitles to align transcripts, and so obtain proposals for the characters in each shot. It was demonstrated on the iconic TV series Buffy the Vampire Slayer. The first of several quirkily-titled papers on this project appeared in BMVC 2006 as “‘Hello! My Name Is... Buffy’—Automatic Naming of Characters in TV Video.” Above and beyond its research contribution, this work popularized the idea of automatically aligning freely available transcripts with subtitles in order to provide weak supervisory information for videos. Mark also developed an algorithm for facial feature detection (corners of eyes, mouth, nose, etc.) using a mixture model of pictorial structures. He generously made the code available, and this detector has in turn been used for projects by many other research groups.

The third project at Oxford turned from recognizing people to recognizing gestures in the form of British Sign Language. Here the supervision, both weak and noisy, was provided by subtitles broadcast simultaneously with the signing on TV programmes. The learning approach used a form of multiple instance learning and was published in CVPR 2009 as “Learning Sign Language by Watching TV (Using Weakly Aligned Subtitles).”

In October 2006, Mark moved to the School of Computing at the University of Leeds, where he continued the theme of reducing the level of supervision by investigating the case of no visual supervision, instead learning models from natural language descriptions alone for recognizing butterfly species. This was an approach for using visual attributes to solve a fine-grained visual categorization task, and was published as “Learning Models for Object Recognition from Natural Language Descriptions” at BMVC 2009. He also investigated how to use the noisy and variable supervision available from sources such as Amazon Mechanical Turk when learning object models. By using multiple instance learning, where the latent variables are the true annotation in the image, he was able both to correct the annotation and learn improved models for human pose estimation. The work was published as “Learning Effective Human Pose Estimation from Inaccurate Annotation” (CVPR 2011). Other contributions included developing novel features that combined local segmentations with the classical HOG descriptor for state-of-the-art object category detection performance (ICCV 2009), and demonstrating the importance of pose-specific appearance models within pictorial structures for human pose estimation (BMVC 2010). This research was carried out together with Mark’s PhD students and postdocs.

What will undoubtably prove to be one of Mark’s major legacies is the PASCAL Visual Object Classes (VOC) challenge to which he was the major contributor. He selflessly devoted so much time to all stages of the challenge: collecting the data, monitoring and checking the annotation, writing the annotation and evaluation software, and overseeing the actual challenge process and workshop. The VOC dataset has been cited in thousands of computer vision and machine learning papers, and hundreds of researchers have entered the annual competitions since their inception in 2005. The challenge suited Mark’s interests in that it objectively and empirically measured performance so that the community could know what really worked. It also enabled innovative new ideas from throughout the community to be explored and disseminated at the VOC workshops. It had been decided—together with Mark—that 2012 would be the final year of the challenge. With Mark’s passing, the final VOC competition and workshop will be dedicated to his memory.

Mark was the quintessential scientist—focused on identifying the essence of a problem in order to understand the general principles underlying it and to make rigorous steps toward a solution. It is clear that even in the all too short time he was with us, he has made the field stronger. We have lost not only a significant researcher, but also someone unboundedly generous in promoting the work of others and in helping colleagues achieve their own potential. He will be greatly missed, but his memory will continue to inspire us all to better things.

There is a tribute website at http://www.bmva.org/obituaries:mark-everingham.