Some Important International Standards and Organizations in Visual Telecommunications

I serien "NORSIGnalet hjelper deg" presenterer vi i dag første av to deler om internasjonal standardisering i visuell kommunikasjon. Andre delen følger i neste utgave.

Av Till Halbach.

Till Halbach har siden 1999 vært stipendiat ved Institutt for teleteknikk ved NTNU. Hans interesseområder er kombinert kilde- og kanalkoding, feilrobust stillbilde- og videokompresjon.

Mail: *halbach@tele.ntnu.no*

The world of international standardization of visual telecommunications is quite complex. I will try in the following to enlighten this area somewhat in an overview; topics to be dealt with are the huge international standardization organizations beyond some of the most important standards made in the area of visual telecommunications, including the standards' history. Technical issues are not discussed. It is the formalities and hierarchies -- how the world works -- on which I will concentrate here.

This overview is split into two parts. I will begin with the world of international standardization organizations, Part I; the second part is then dedicated to the standards. But be aware, standards people love abbreviations; however, I will explain all of them. It is one of the intentions of this article to make the standardization world more understandable.

Part I: Organizations in international standardization

There are basically two types of players: Profit and non-profit ones, some also say non-governmental and governmental ones, respectively. Profit organizations design a standard to be able to sell it successfully at more or less high fees to their customers later. In this case, it is also said that there are royalties on the standard; you pay actually for the right to use a patented technique or algorithm (also called IPR, Intellectual Property Rights). On the other hand, non-profit organizations are mainly interested in such noble issues like technical proceeding of mankind and intercommunication among all countries of the world. They aim to produce royalty-free standards which can be acquired to only small fees which cover printing, delivery, and administration.

Two large organizations that collaborate tightly for standardization efforts mainly form the non-governmental group in visual telecommunications: 7International Organization for Standardization (ISO) and International Electrotechnical Commission (IEC). ISO was founded in 1947; its mission is "(...) to promote the development of standardization and related activities in (...) scientific, technological and economic activity". The IEC can look back on a history beginning in 1906 and "(...) prepares and publishes international standards for all electrical, electronic and related technologies". Both organizations consist mainly of middle-sized to large companies, but also educational institutions like universities have joined. A country can become member of ISO/

IEC through formation of a so-called National Body (NB). Standardization work is carried out by participating NBs in priorly established technical committees; they address only well specified and highly specialized technical fields. Norway is usually represented via Norges Standardiseringsforbund, but -- compared with giants like USA or Germany -- only to a small extend.

There is an almost uncountable number of small groups under the ISO/IEC body. I will therefore concentrate myself on just one organ: The Joint Technical Committee (JTC) with the number 1, JTC 1, which has been established for dealing for information technology issues; see also Figure 1 which explains group relationships. Domiciled in JTC 1 is, among others, Subcommittee SC 29 (the SC with the number 29), owing the "name" Coding of Audio, Picture, Multimedia and Hypermedia Information. SC 29 includes the two important working groups WG 1, entitled Coding of Still Pictures, and WG 11, Coding of Moving Pictures and Audio. We have hereby reached the bottom of the hierarchy pyramid.

ISO/IEC JTC 1/SC 29/WG 11 is the formal name of, and identical with, MPEG. Many say the M stands for Motion but, in fact, it is short for Moving. So, MPEG is Moving Pictures Experts Group. MPEG was formed in 1988 to establish standards for coding of both video and associated audio. Participation in MPEG is popular -- usually more than 300 experts from some 200 companies and other institutions take part in at least three annual meetings; about 20 countries (including Norway) are on the average represented. This results not unusually in more than 500 contributions/documents per meeting! So far, the standardization group has worked out the three famous standards of the same name, so don't get confused now, MPEG-1, MPEG-2, and MPEG-4. Current projects are, see also NORSIGnalet Nr.1 (March 2001), MPEG-7, and MPEG-21. This leads to the interesting problem of predicting the number of the next MPEG standard. (The problem is left for the reader; its solution is -34! NORSIGnalet can present the solution steps in its next number in case of interest.)

In order for the reader to get a useful impression of a normal standardization procedure, I will describe it in more detail in the following. The standardization process begins with a basic algorithm which is typically evaluated through a CfP, Call for Proposals, and a subsequent competition of the proposals handed in. This model is then, through adoption of meeting proposals/contributions, extended, limited or refined. Often, subgroups named

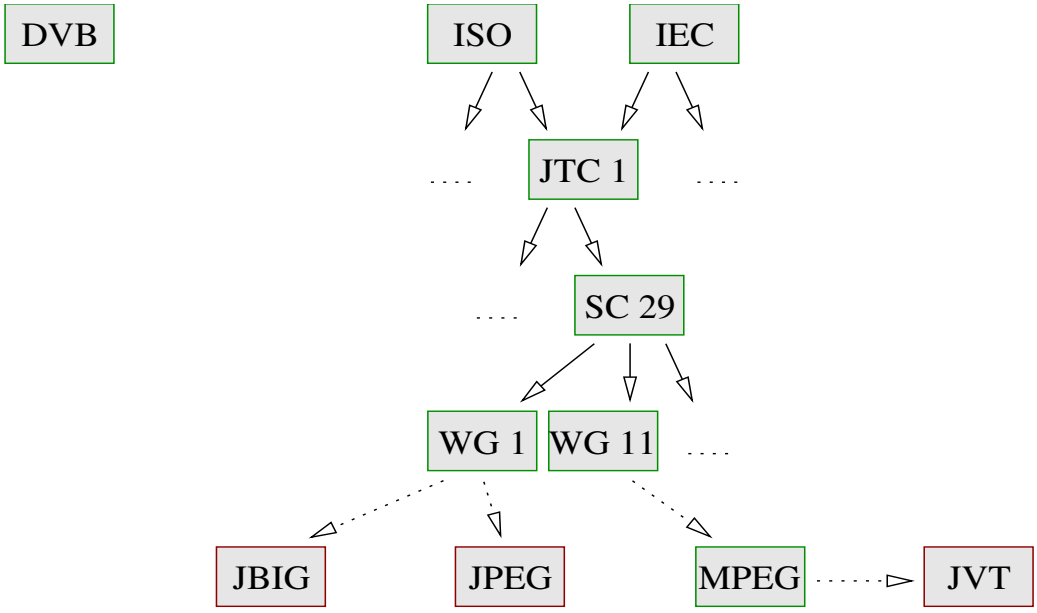

Figure 1: Non-govenmental organizations. Solid-line arrows denote hierarchical dependencies, and dotted lines show identities.

Ad Hoc Groups (AHGs) are formed that concentrate on only one certain topic. These groups carry out so-called CEs, Core Experiments or simulations, that in turn investigate a specific problem. The decision process for ISO/IEC standards runs usually through several distinct stages, through which the later standard changes its status; from WD (Working Draft), later CD (Committee Draft) and Final Committee Draft (FCD) to Draft International Standard (DIS), and terminating in the final document IS (International Standard). Most document changes are made before reaching the status WD, and document changes from e.g. DIS to IS are very small. So when citing a standard, it is obviously of advantage to have an IS as reference. A becoming standard in DIS status has to be balloted on a yes/no basis by participating NBs and is accepted as IS if at least 75% of the NBs cast a positive vote. I would finally like to mention that, sometimes (example JPEG2000), FIS documents are freely available, whereas a fee has always to be paid for IS.

On the still-image compression front, there are other organs. As mentioned before, there is another WG of SC 29 (you don't have to look up abbreviations now, do you?), WG 1, which is identical with the two communities JPEG (some say P is short for Pictures, but that's only valid for MPEG, so it actually stands for Photographic, hence Joint Photographic Experts Group), and JBIG (Joint Bi-level Image Experts Group). The "Joint" in the name refers to a collaboration with ITU-T. Even though JPEG, founded in 1986, convened officially under ISO's rule, it was closely coordinated with ITU's CCITT SG VIII (abbreviations explained later). JPEG worked so far mainly on the famous still-image compression standard named after the group, so -- again -- don't get confused. Current efforts have names like JPEG2000, JPEG-LS and Motion JPEG2000. JBIG on the other side worked, among other standards, on JBIG1, and is currently busy with producing JBIG2. WG 1 nowadays has over 100 active members from about 20 countries (including Norway), and conducts three main and some interim meetings per year. The standardization

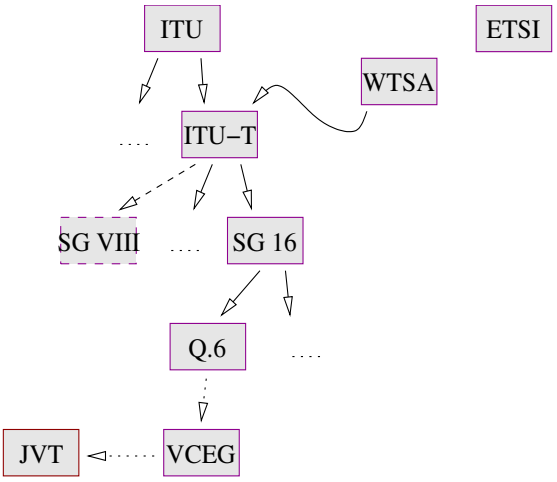

Figure 2: Governmental organizations. Solid-line arrows denote hierarchical dependencies, dotted lines mean identities, and dashed lines show former structures.

process is identical with the one described for the MPEG-x standard series.

In case you are not bored to death yet, let's turn over to the governmental organizations and their most important player: The ITU, the International Telecommunication Union. ITU is defined as "(...) the United Nations' specialized agency in the field of telecommunications". It was formed in 1865, at that time named International Telegraph Union. So, it has quite a history. Today, ITU is a huge organization with many committees, subgroups, and other organs, see also Figure 2. Of special interest for us is the ITU-T, the Telecommunication Standardization Sector of ITU, a permanent organ dealing with telecommunication standardization issues. It was created in 1993, replacing the CCITT (International Telephone and Telegraph Consultative Committee), a former committee within the ITU.

So-called Study Groups (SGs) carry out the standardization efforts; an important one being SG 16, labeled Multimedia Services, Systems and Terminals. Topics for study by ITU-T's SGs are decided on by the World Telecommunication Standardization Assembly (WTSA), a large symposium of more than 600 delegates representing at least 60 countries. The WTSA meets once every four years and consists of ITU members. As you see, we have already reached a high degree of structural complexity.

Another group under SG 16, named Q.6 (Question 6), is known as Video Coding Experts Group or short VCEG (the full name is ITU-T Q.6/SG 16). VCEG was formed in 1997 for the development of further advancement of video standardization techniques. This explains also its label AVC, Advanced Video Coding. Addressed issues have been the development and maintenance of the video coding standards H.120, H.261, H.262, H.263, H.263+, and H.263++. Products worked out by VCEG are so-called Recommendations which in general correspond to a specific IS of ISO/IEC, or vice versa, something which happens actually more often. I.e. an IS may adopt a previously released Recommendation. More on that later in Part II. In fact, ISO/IEC and ITU-T often collaborate in fields of mutual interest. Recently, in 2001, one further collaboration saw the light of the day: The Joint Video Team (JVT). This group was established to bale the developments of both MPEG and VCEG, see Figure 3 and 4. The effect is that standardization efforts be concentrated and harmonized; more, it will lead to two new standards which actually are one: H.264 or MPEG-4 Part 10 (AVC), at the time being also known as H.26L. VCEG and JVT meet about 5 times a year. Decisions are made by consensus under the normal development process, i.e. there are no stages such like WD or FCD; later, for approval of a Recommendation, a positive 2/3 majority of its balloting members is required. VCEG on its own consists of about 70 members, whereas participation in JVT is with some 180 persons quite strong; about 30 countries are represented.

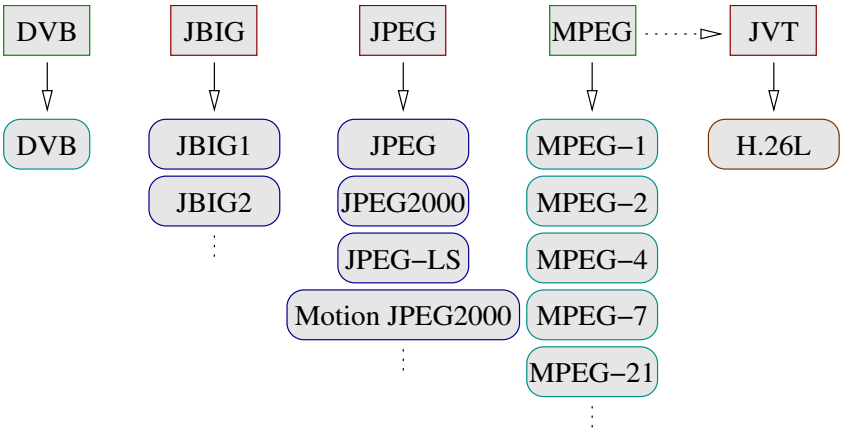

Figure 3: Important still-image and video coding standards. Solid-line arrows symbolize project responsibilities, and dashed lines mean identities.

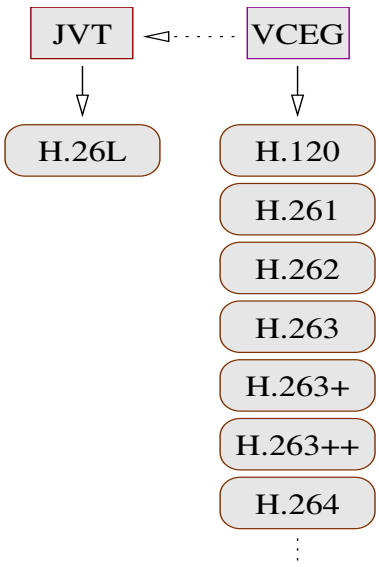

Figure 4: ITU's video coding standards.

For complementing purposes, I also want to mention the DVB project which is led by a non-governmental industry consortium with the same name. DVB stands for Digital Video Broadcasting and refers to a bunch of global standards for delivery of digital television and data services, mainly addressing visual broadcasting matters. It consists of over 300 members representing more than 35 countries. The DVB standard is mainly based on other standards like MPEG-2 with a certain set of own parameter settings and other specifications. Since DVB is no official standardization institution, the final specifications are then offered for standardization to relevant organizations like ETSI, ITU-T, and others. In case I have forgotten to write about ETSI: It is the European Telecommunications Standards Institute that unites over 90 members from more than 50 countries inside and outside Europe. In addition to the world-wide standardization associations, ETSI must be seen as a governmental organization to strengthen European needs and

interest in world-wide standards, and it works tightly with organizations like ISO/IEC and ITU-T.

Alright, we are through. One final note: Organizations are changing; names alter internally, new committees are founded, others abandoned, collaborations are established or finished. The description here can therefore only be a contemporary mirror. If you are interested (I assume so; otherwise you would not have read until here) I recommend to follow the press releases of ITU-T, ETSI, and ISO/IEC to update yourself.

Websites for further information about the organizations (*http://*): *www.iso.org, www.iec.ch, www.sc29.org, www.jpeg.org, mpeg.telecomitalialab.com, www.itu.int, www.itu.int/newsarchive/wtsa2000* (from the archive), *www.etsi.org, kbs.cs.tu-berlin.de/~stewe/vceg* (slightly out of date), *www.dvb.org,* and *www.standard.no.*

Other useful societies (*http://*): *vqeg.org* (Group for video quality assessment), *smpte.org* (film, television, video and multimedia) and *www.imtc.org* (standards promoting and proposals).

Fors. fra side 18

det forventede vokabularet, med de tilhørende regionale uttalevariasjoner, er også viktig for å oppnå god talegjenkjenningsytelse.

Tilgjengelige språkressurser varierer fra språk til språk. De er best utviklet for engelsk, og et fåtall andre store språk. Mangelen på gode språkressurser er faktisk et av de største problemene knyttet til utvikling av tjenester basert på automatisk talegjenkjenning for norsk.

3 Applikasjonsdomene

Vi har sett at det er mange aspekter ved et språk som kan påvirke ytelsen til automatisk talegjenkjenning. Det er imidlertid applikasjonsdomenet som er helt avgjørende for betydningen av de forskjellige aspektene for talegjenkjenningsytelse. For eksempel, hvis talegjenkjenning skal brukes til å gjenkjenne pinkoder, er det helt uinteressant om det finnes mange bøyningsformer i språket. Det er like uinteressant om det finnes ord som skilles bare i tonelaget, så frem ingen av sifrene skilles bare i tonelaget. Det som er interessant er hvor forskjellige sifrene er fra hverandre.

Applikasjonsdomenet bestemmer både vokabularet som skal inkluderes i uttaleleksikonet, og språkmodellen som vil bli brukt. Derfor vil enkelte språk som er lett å gjenkjenne for et applikasjonsdomene vil kunne være vanskelige for et annet applikasjonsdomene. Et eksempel på dette er gjenkjenning av telefonnummer som er vanskeligere for norsk enn for mange andre språk siden vi bruker to forskjellige tellesystemer. Tallet 67, for eksempel, kan bli uttalt både som *sekstisju* og *sjuogseksti*. I tillegg kan tallene 7, 20, 30 og 40 uttales på to forskjellige måter hver. Derfor vil både språkmodellen og utaleleksikonet være mer komplekse for norsk enn for mange andre språk for dette applikasjonsdomenet.