



## OPTIMIZING VIDEO CLIPS IN EDUCATIONAL MATERIALS

Aleš Oujezdský

Department of Information and Communication Technologies, Pedagogical Faculty,  
University of Ostrava, Fráni Šrámka 3, Ostrava-Mariánské Hory, Czech Republic  
ales.oujezdsky@osu.cz

### Abstract

The use of videos from digital camcorders has become a standard in education in recent years. The curriculum is easily accessible and appeals to a wider audience. The lessons use videos of various physical processes and chemical experiments. However there can be problems with this format. The video quality is often degraded in the final stage when the video is being prepared for placement in education. These include teaching materials in the form of web pages, e-learning courses or flash multimedia objects. The final product of editing video from a digital camcorder is a DVD video. However, if we want to transfer this to the Web or other educational material, it is necessary to remove non-square pixels, interlaced video and choose the appropriate compression. For these operations, there are many interpolation algorithms (nearest neighbour, bilinear interpolation, bicubic interpolation), filter deinterlacing (wave, bob, blend), and compression tools. By selecting appropriate settings for these parameters, the video material can be optimized while maintaining the highest possible image quality. The final step before publishing the video is its conversion into one of the used codecs. Codec's settings will largely impact the final quality and size of the video-clip.

### Keywords

optimizing video (resolution, interpolation, interlace, deinterlace, compression)

### Introduction

Videos have become an integral part of educational materials, courses and multimedia objects. They are found in e-learning courses, they are part of courses in LMSs. The advantage of videos is their clarity. If they are accompanied by an audio track, or even interactive elements, they can better explain the teaching substance and they affect additional senses. Video illustration is much more effective than the same problem being described by text and images.

Video can be obtained in several ways. One method of creating educational videos is screen capturing; when you scan activity on the computer screen and store everything in a video file. This way we can create effective video tutorials for working with various software applications, or describe other computer activities. Some applications for screen capturing allow you to insert text captions and buttons into an already scanned video. This way the video is enriched with interactivity, learners themselves control individual steps of the video. The advantage of these videos is that they are usually saved into a suitable video format, without interlacing.

Another option is to obtain a video recording using digital camera. This way, we can create our own videos of various physical processes and chemical experiments, or even a documentary video describing various places which students couldn't visit themselves. Video is saved into one of the standard formats that are used by digital cameras.

However, if we want to use this video on the web, it must be optimized, especially in terms of resolution, interlacing and compression. To do this we need to learn some methods and procedures. Although optimization methods are demanding on computer's processing power, current desktops and laptops have sufficient power to process them. The optimization can even be done on low performance machines, with the duration of optimization being inversely proportional to the computer's power.

## Optimization Methods

Digital video recorded using a digital camera in standard definition (SD) has parameters in accordance with the PAL system that was used for analogue television signals. Nowadays, this video has resolution of 720x576 pixels, frame rate of 25 frames per second, the ratio of 4:3 or 16:9 at the same resolution through non-square pixels and the video is interlaced.

It is the non-square pixels what causes problems during the conversion. The cameras record video in the resolution of 720x576 pixels, both in the ratio of 4:3 and 16:9. However, this resolution does not correspond to either ratio. If we keep the video at this resolution and do not address the aspect ratio, it will result in a distorted video and especially the aspect ratio of 16:9 would be a remarkable distortion.

In an interlaced video, the image is stored in individual half-frames, where the entire frame consists of two fields. The first field contains the odd lines of the overall frame, the second field contains the even lines of the total frame. Computer monitors, however, display video in full frames (progressive), and when you play interlaced video, two fields are displayed simultaneously. This is not so apparent in slow scenes, because the difference between the two fields is 20 ms. In faster scenes however, there are the so-called comb artefacts that need to be removed.

The last step in the video optimization is video compression with a suitable, correctly adjusted codec. One of the important conditions when choosing the codec is its availability and usability in various devices. The aim of compression is to create a video file that is as small as possible with the highest possible quality. It cannot be generalised that the smaller the video file is the worse the quality. Current modern codecs use sophisticated algorithms for video compression, which allow to create small video files that retain high quality of the image.

Several software products are available for video optimization. Video editing software used for editing digital camera videos does not allow our desired setting of the output format. It is therefore desirable to just export the video with the lowest possible compression without losing quality and use another application for the optimization. Virtual Dub is an example of a suitable application. It is a freeware that can not only be used for compressing, but it also handles basic editing, use of a series of filters for resolution change, deinterlacing and video compression.

## Optimizing Resolution

If you want to change the resolution of a video, you must first calculate the required resolution in the correct aspect ratio in order for it to be displayed correctly, as shown in Figure 1.



**Fig. 1:** Incorrect video resolution on the left, resolution after conversion on the right

Video from a camera has the resolution of 720x576 pixels. To have the video displayed correctly after conversion into appropriate format, it is also necessary to convert the resolution. After conversion, the final resolution is as follows:

- for the ratio of 4:3 it is 768x576 pixels;
- for the ratio of 16:9 it is 1024x576 pixels.

The use of the interpolation method is a very important setting element that plays a significant role in the final quality of the recalculated video. It is necessary to use interpolation because by changing resolution the previously non-existent pixels have to be added. Interpolation is actually finding the approximate (most accurate) values of the RGB pixel that is artificially created. There is a wide range of interpolation methods and they are used in the processing of digital photos (reducing or enlarging images), in processing digital video, and also in a large variety of mathematical problems.

Interpolation methods are classified as non-adaptive and adaptive. Non-adaptive methods process the entire image regardless of its content. Adaptive methods attempt to understand the character of different areas in the image and choose the most suitable interpolation methods. When converting a video, these types of interpolation are used most often:

- Nearest Neighbour – is the simplest interpolation method, which creates a new pixel by copying the neighbouring pixel. This method retains the sharp edges, but the image is aliasing (jagged edges). This method is least computationally demanding.
- Bilinear interpolation – this interpolation method is based on the fact that the new pixel is created by a weighted average of the surrounding 4 pixels (2x2 matrix). This method suppresses aliasing but the final image is blurry and needs to be sharpened. This method is computationally more demanding than the nearest neighbour.
- Bicubic interpolation – is based on the fact that the new pixel is created from the surrounding 16 pixels (4x4 matrix); when calculating the weighted average, closer pixels weigh more than distant pixels. The method does not produce aliasing and produces very smooth transitions. The image must also be sharpened. The disadvantage is a large computational complexity.

## Deinterlace

Interlaced video was originally used for analogue broadcasting and was created to save half the bandwidth by transmitting only half of the information. It did remain to be used in TV however. Even today's digitally broadcasted televisions still use interlaced video. LCD and plasma TVs

display video image progressively, and therefore they must include filters to remove comb artefacts of an interlaced video. Figure 2 shows how interlaced video is displayed on the computer screen. Static objects (in the background) are fine and sharp. With moving objects and dynamic scenes however, visible deformation occurs due to interlacing.



**Fig. 2:** Interlaced video

While watching a DVD movie with interlaced video on a computer, software players automatically filter interlacing in real-time. The problem occurs while converting the video into a different format such as MPEG4 in order to place this video on a web. This video needs to be much smaller but its quality has to be maintained. In this case, interlacing would be saved into the final video file and it would not be possible to remove it with filters while playing. Therefore it is necessary to deinterlace the video before such a conversion. There are many methods used for deinterlacing:

- Weave – a method based on including an alternation of even and odd lines from both fields into one frame. There is no image conversion and therefore this is a very fast method. It is mainly used for static scenes with little movement. There are significant comb artefacts.
- Blend – a method that uses the averaging of individual fields. Both fields are put over each other and the pixel values are averaged. There are no comb artefacts but „ghosting“ is formed.
- Duplicate – a method that uses only one field, the second one is dropped. There is a loss of continuity because by dropping one field, half the time information is lost.
- Bob – a method when both fields are duplicated in order to avoid half time information. However, there is an artefact of hopping that is caused by duplicating the lines in both fields.
- Adaptive deinterlace – a method based on comparing two fields or frames and looking for changes in the image. Static scenes use weaving, dynamic scenes use blending. This method is fast and has good results.

- Motion compensation – the most advanced method, but also computationally most demanding. This method watches the dynamic scenes between two fields and tries to compensate for them in the entire frame.

Various applications for video editing and conversion use various methods for video deinterlacing. Often it is impossible to choose the used method or to set other parameters, such as interpolation method used for deinterlacing and calculating pixels.

## Compression

The size of an uncompressed video is enormous and it would be difficult to place such a video on the web. It needs to be compressed. Codec is used for video compression. It is a blend of two words – coder and decoder. Coder takes care of the video compression while decoder decompresses the video when it is played.

There are numerous codecs in existence and every video is compressed in one or other way. The used codec and its setting will greatly impact the size and quantity of the final video. A poorly chosen codec can completely degrade the resulting video quality. The compression can be lossy or lossless, similarly to graphic and audio formats.

From the additional video usage, codecs can be divided into two categories. The first category consists of codecs that are used with videos designated for further editing. Their aim is to lower the size of the video while retaining the best possible picture quality. All lossless codecs and the DV codec used by DV cameras with magnetic tapes belong to this category, as well as the MJPEG codec (compressing each frame of the video separately). The second category is made of codecs designated for video finalization. These videos are then used on the web, or saved on DVDs and BDs or put into learning multimedia objects. The picture quality is close to the original, mild picture degradation does not matter as it is imperceptible by naked eye. All MPEG codecs belong to this category as well as the DivX and Xvid codecs. Further editing of these video is problematic, however, because the video is made of key frames, and the remaining frames are computed of them.

One of the most used is the Xvid codec that follows the MPEG4 ASP standard. This codec uses frequency conversion for compression, with the highest coefficients being deleted. The picture is divided into smaller blocs, usually the matrix of 8×8 points. These blocs are then converted using discrete cosine transformation (DCT). Similar principle is used with the JPEG format.

The DCT compression method alone is not sufficient, however, and additional methods and complex algorithms are used for bigger compression while retaining the picture quality. The main idea is that individual frames of the video are very similar. That is why all that needs to be encoded is the change between individual frames and not the frames themselves. Obviously, this method has its disadvantages. Encoding takes longer times and several previous frames need to be decompressed simultaneously to calculate the frame being currently played. The codec uses the following three types of frames:

- I (intra coded) the frame is compressed independently and does not require other frames for its display.
- P (forward predicted) the frame requires one previous frame of the I or P type.
- B (forward and backward predicted) these frames are based on the previous as well as the following I or P frame.

The reasoning behind the whole algorithm is the assumption of motion picture. The whole group of frames is called Group of Pictures (GOP). While decompressing a specific P-frame, we need

to find the previous I-frame and then decompress all the following P-frames until we hit our specific P frame. With B-frames both adjacent P-frames need to be decoded. The sequence and interdependence of individual frames is depicted in the Figure 3.

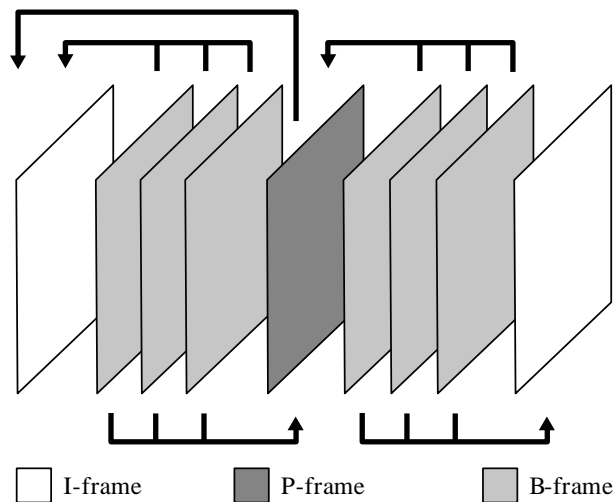


Fig. 3: Sequence and interdependence of individual frames

When configuring the codec, there is a large number of settings that can be used to optimize the resulting size and quality of a video. These are the most important parameters:

- Target bitrate – enables to set the bitrate of the video. Another option is to set the required size of the video and the codec will automatically calculate appropriate bitrate.
- Encoding type – enables one- or two-way encoding. Two-way encoding has better results, but takes significantly longer.
- Motion – choice affecting the subsystem for motion evaluation. The higher the value, the more precise the evaluation will be. Better evaluation saves more bits.
- VHQ – higher VHQ value enables to decrease the quantisation value and increase the quality. The compression will be significantly higher, but the encoding time will rise dramatically. Low value (1) should be used for high bitrates, and higher values (2-4) should be used for low bitrates.
- Cartoon Mode – choice for better encoding of cartoon movies.
- Greyscale – choice for better encoding of black and white movies.
- Max B-frames – this setting determines how many B-frames can be used in sequence. Generally speaking, values between 0 and 2 are suitable for „natural“ movies and higher values are used for animated movies with a high volume of single-coloured surfaces.
- Quarterpel – Xvid uses half-pixel accuracy for motion detection but can use quarter-pixel accuracy. The idea is to find better motion vectors to reduce the bitrate while retaining picture quality.
- GMC – methods helping with objects that are zooming in or out, or are otherwise in motion without actually changing. This covers frequent zooming or handheld camera work. Applies mostly to amateur films from holidays, etc.

Options for Xvid codec settings are shown in the Figure 4. User can select from many parameters or let the codec calculate the bitrate based on the requirements on the video size.

Standard bitrate of the Xvid codec is around 1000 kbps. The issue of proper codec setting is much more complicated and complex however and the results are often based on empirical methods.

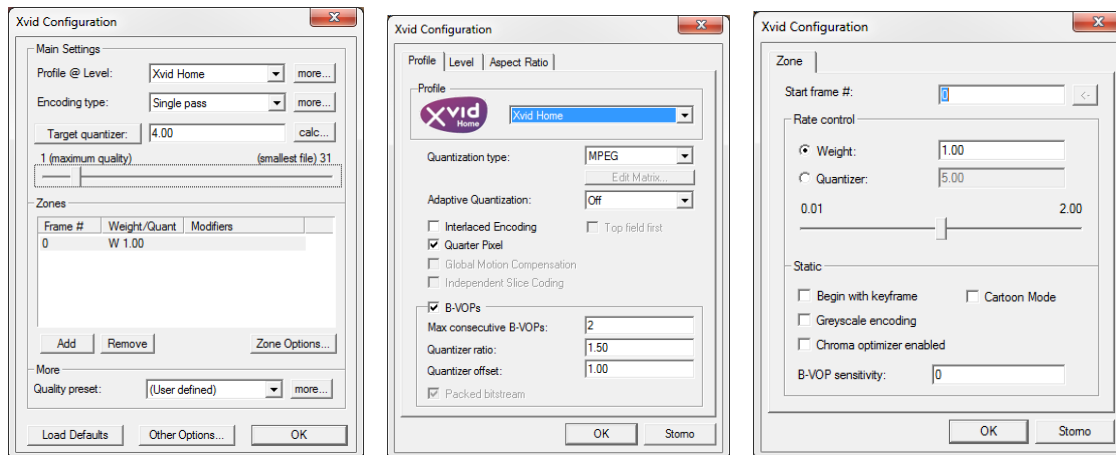


Fig. 4: Xvid codec settings

## Optimization Results

Several methods for conversion of video resolution and deinterlacing can be tested. For example, in Virtual Dub, video from a digital camera Panasonic SDR-H250 HDD is a source material. Video resolution is 720x576 pixels, frame rate is 25 frames per second, the aspect ratio is 16:9 and the video is interlaced.

Virtual Dub also enables compression in a codec of choice, in our case the Xvid codec.

## The Change of Resolution

Three methods of interpolation have been chosen to change the resolution of the original video: nearest neighbour, bilinear and bicubic interpolation. Video has been optimized in the Virtual Dub application. Resize filter setting is shown in Figure 5.

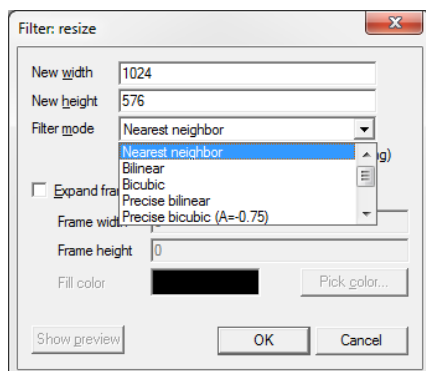
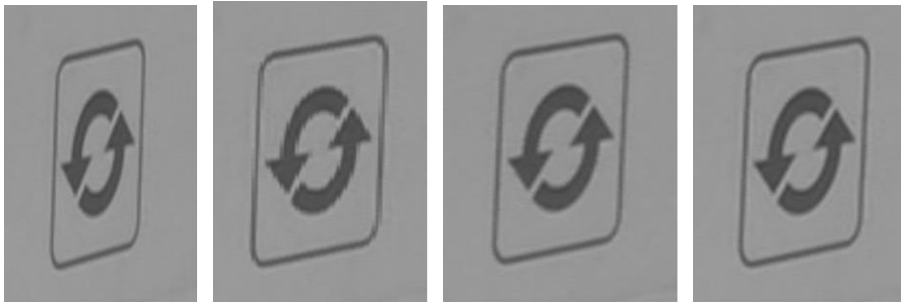


Fig. 5: Basic choice of interpolation method in Virtual Dub

The results of the interpolation methods are shown in Figure 6. It is a detail cut from the same video frame. The original video, with the incorrect aspect ratio, is on the left. The following is



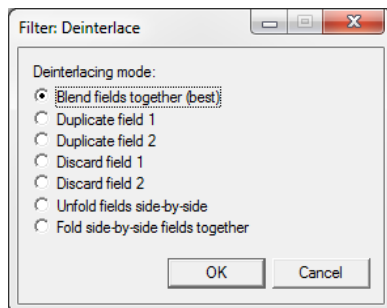
the result of the nearest neighbour interpolation method, bilinear interpolation and bicubic interpolation on the right.



**Fig. 6:** From left: Original video, nearest neighbour, bilinear, bicubic interpolation.

### Deinterlace

Four methods have been chosen for deinterlacing: Weave, Blend, Duplicate, and Smart deinterlacing, which is an adaptive method with motion detection.



**Fig. 7:** Basic choice of deinterlacing method in Virtual Dub

The result of the Weave method is shown in Figure 8. There are odd and even fields in one frame, and the time lag of 20 ms between the fields is significantly noticeable. Comb artefacts completely degrade the quality of video playback on the computer.



**Fig. 8:** Weave Method

Figure 9 shows the output from the Blend method. Comb artefacts were removed by averaging the two fields, but the so called „ghosts“ have formed.





**Fig. 9:** Blend Method

Duplicate method is shown in Figure 10. Odd frame is duplicated, even frame is dropped. Comb artefacts are noticeable, edges are jagged.



**Fig. 10:** Duplicate Method

In our case, the best deinterlace method is Smart deinterlace. The result of this method is shown in Figure 11. Although this method is demanding on computer performance, it brings the best results.



**Fig. 11:** Smart Deinterlace Method

## Compression

Final video compression has a significant impact on the size of the final video file but also on the image duality of the video. Modern codecs are able to compress video into size that is 4× smaller than a DVD video, while retaining the same quality.

As a testing input video we used a 10 second video clip that was deinterlaced with Smart Deinterlace filter and the original resolution of 720×576 was converted to correct 768×432, which has the aspect ratio of 16:9. The video intentionally has both static and dynamic scenes. It comprises of static background and a passing cyclist and bus. No sound is present and therefore no compressed audio track. Figure 12 shows image from the clip.



**Fig. 12:** Frame from the compressed video

Several different compressions were used for the 10 s video:

- MPEG2 – original codec used for saving the video in the camera. Bitrate is 5000 kbps. The video file size is 5.9 MB. The quality is fair, the sunny weather provided enough lighting, there is no visible noise in the video.
- Uncompressed format – video was exported without compression. Bitrate is 200 000 kbps. The video file size is 237.27 MB. Picture quality is the same as the original. Twitches occurred during playback due to the enormous data rate.
- Xvid codec – video was exported with a 2000 kbps bitrate. Default setting of the codec. Resulting file size is 2.4 MB. Picture quality is identical with the original recording.
- Xvid codec – video was exported with a 1500 kbps bitrate. Default setting of the codec. Resulting file size is 1.6 MB. Picture quality is almost identical with the original recording.
- Xvid codec – video was exported with a 1000 kbps bitrate. Default setting of the codec. Resulting file size is 1.1 MB. Picture quality is sufficient in comparison with the original recording.
- Xvid codec – video was exported with a 700 kbps bitrate. Default setting of the codec. Resulting file size is 0.9 MB. Picture quality deteriorated a little in comparison with the original recording. Colour maps are visible with pixelisation of the static background. Dynamic scenes are blurry.
- Xvid codec – video was exported with a 500 kbps bitrate. Default setting of the codec. Resulting file size is 0.6 MB. Picture quality deteriorated significantly in comparison with the original recording. Large colour maps are noticeable with visible colour transitions and significant pixelisation of the static background. Dynamic scenes are blurry.
- Xvid codec – video was exported with a 2300 kbps bitrate. Motion setting has been turned off. Codec does not use prediction and motion detection change. B- and P-frames are not used, only I frames. Resulting file size is 3 MB. Picture quality is poor, the results are worse than those of the 500 kbps default setting Xvid encoding.

## Optimization Assessment

The results of the video optimization show that it is necessary to use filters for conversion. It is also important to make sure the correct filter is chosen because the results of different methods are different. The computing power of modern computers is sufficient enough that we do not have to take into account the load on the computer and the time needed for video conversion.

Bicubic interpolation is the best choice for resolution changes in relation to aspect ratio. Bilinear interpolation achieves good results in less time and is fully sufficient for some videos. The „Nearest Neighbour“ interpolation method is indeed the fastest but produces visible aliasing and it is therefore not suitable. Since the bilinear and bicubic transformation create subtle blur of the image, the video can be sharpened using a filter.

Weave method is the least desirable for deinterlacing dynamic scenes. This method only combines even and odd fields but does not deal with comb artefacts. Blend method creates the „ghosts“ effects. Duplicate method brings better results than Weave but significant comb artefacts are visible in the edges. Smart deinterlace method is the best. It respects the scene dynamics and it is possible to set one's own interpolation method to calculate the pixels.

The main role for video optimization is played by the final compression in the chosen codec. MPEG2 is not suitable for web-used videos; MPEG4 is a better alternative. For a standard definition (SD) video, the Xvid codec, belonging to the MPEG4 ASP family, is a good choice. It offers good quality and small video file size. Default setting of the codec with the choice of a 1000 kbps bit rate is sufficient for the compression. The codec uses Motion Detection function and also I-frames (intra coded), P-frames (forward predicted) and B-frames (forward and backward predicted). If the Motion Detection is turned off, the video quality is unusable. For different types of movies we can utilise additional functions to improve the picture quality, such as *GMC* for handheld camera movies, *Cartoon Mode* for cartoon movies, or *Greyscale* for better encoding of black and white movies.

## Conclusion

Students at the Department of Information and Communication Technologies learn about the processing of digital videos and apply this knowledge in the creation of teaching materials.

It is necessary to optimize videos from digital cameras in standard definition (SD) that are to be used for educational materials on web sites and in the LMS as a part of e-learning courses or as a part of multimedia objects. Of particular importance is the conversion of video resolution with regard to aspect ratio and video deinterlace. There are many interpolation methods that are used for conversion of video resolution, and deinterlacing filters that are used to deinterlace video. When choosing the appropriate method, it is not efficient to only take the speed of the conversion into consideration, as the fastest method results in the worst quality. It is best to use adaptive methods that take the dynamics of changes in individual video frames into consideration.

Choice of the codec and compression methods plays a key role. Suitable setting of the codec will ensure the smallest possible size of the video file with optimum picture quality.

A good alternative is to use a camera with progressive video storage which in addition uses square pixels. Some digital cameras currently have high-quality video recording and that is usually a progressive video with correct resolution to the aspect ratio.

## References

MEŇHART, Pavel. *Odstranění prokladu v nekomprimovaných digitálních videosekvencích: Deinterlacing in uncompressed digital video sequences*. Brno: Vysoké učení technické, Fakulta elektrotechniky a komunikačních technologií, 2009. 1 elektronický optický disk [CD-ROM / DVD]. Diploma thesis.

SLAVÍČEK, T. Všudypřítomná interpolace. *DIGIfoto: největší časopis o digitální fotografii*. Brno: Computer Press Media, 2005, č. 7, s. 14-17. ISSN 1801-0873.

OUJEZDSKÝ, Aleš a David GROBELNÝ. Digital Video Processing. In: *Information and Communication Technology in Education: Proceedings of the 2008 ICTE Annual Conference*. Ostrava: University of Ostrava, 2008, s. 128-131. ISBN 978-80-7368-577-5.

BIJNENS, Marie, Mathy VANBUEL, Soetkin VERSTEGEN a Clive YOUNG. Handbook on Digital Video and Audio in Education. *VideoAktiv* [online]. 2006 [cit. 2012-11-13]. Dostupné z: <http://www.videoaktiv.org/>