Simon Fraser University Surrey

School of Interactive Arts and Technology

Wide-Span MCI*†

David Steffen Bergman, Belgacem Ben Youssef, Jim Bizzocchi

{dbergman, bbenyous, jimbiz} @sfu.ca

Motion Compensated Interpolation (MCI) uses Motion Estimation (ME) followed by Motion Interpolation (MI) to create new frames for frame-rate up-sampling and slow motion video. Both ME and MI are fundamental problems in digital video processing and have been the subject of much research. Furthermore, image-plane motion estimation is an integral part of motion compensated filtering and compression (eg. MPEG-2).

We know from Physics that an object in motion tends to stay in motion. This means that linear motion is more likely than oscillatory motion, and vectors that span a few frames may be better able to depict object motion. Although pixel-wise MCI was abandoned more than a decade ago, recent hardware advances make it an emerging viable alternative to conventional region-based techniques.

A small amount of motion can result in large differences in the pixel values of a scene - especially near the edges of an object. Also, a wider field of view and higher resolution (e.g. HDTV) makes it harder to fool the eye, and therefore more difficult to do motion compensation. The effectiveness of any MCI algorithm will depend not only on the ME and MI phases, but also on source video colour precision and scene complexity.

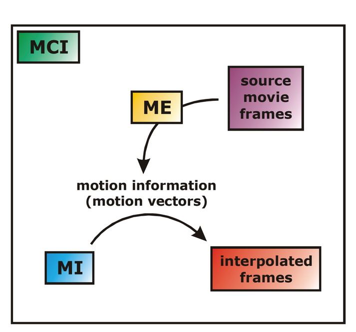

Figure 1. Motion Compensated Interpolation (MCI) is a process that combines Motion Estimation (ME) and Motion Interpolation (MI). Motion estimation looks across frames to see how sum-of-luminance (SOL) values migrate. It uses a 5-level hierarchy of down-sampled luminance images for each source frame in the video sequence. After ME, the interpolation step uses the resulting motion information to produce new video frames.

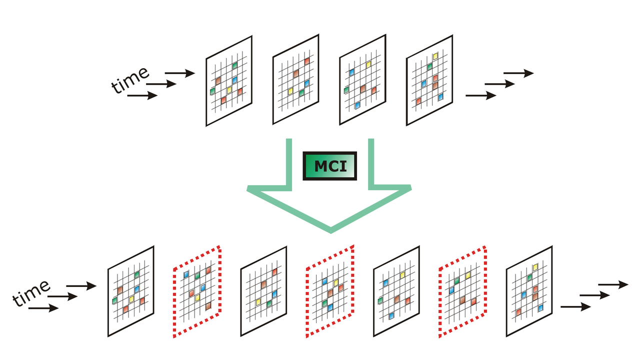

Figure 2. The goal of Motion Compensated Interpolation (MCI) is to produce new video frames that look as though they belong within the video sequence. The MCI process above shows only one new MCI frame inserted into the video sequence per each source frame time gap (new frames are in red). Note that there can be any number of interpolated frames produced between source frames.

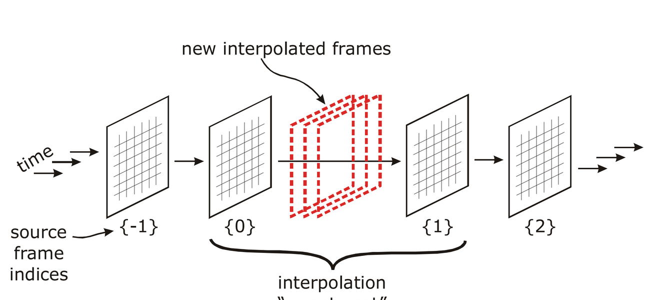

Figure 3. Overview of MCI showing source video frames and the interpolation "sweet spot" where the new interpolation frames will be placed for this MCI iteration. A movie is composed of discrete video frames where each frame is a 2D image. MCI can be iterated for each time gap in the source video sequence to produce a video sequence with new frames. This results in slow-motion if the movie's frame rate is preserved.

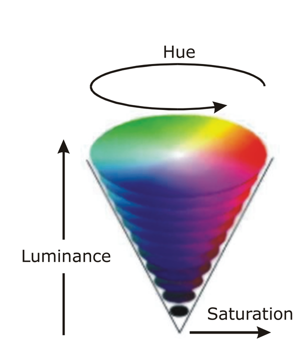

Figure 4. The HSB (Hue, Saturation, Brightness) colour model shows that luminance values (the brightness) are the same throughout each horizontal cross-section of the HSB colour-space. In other words, a red can have the same luminance value as a blue, green, and orange (etc.) as long as the colours' positions in the HSB model are on the same hypothetical luminance disk.

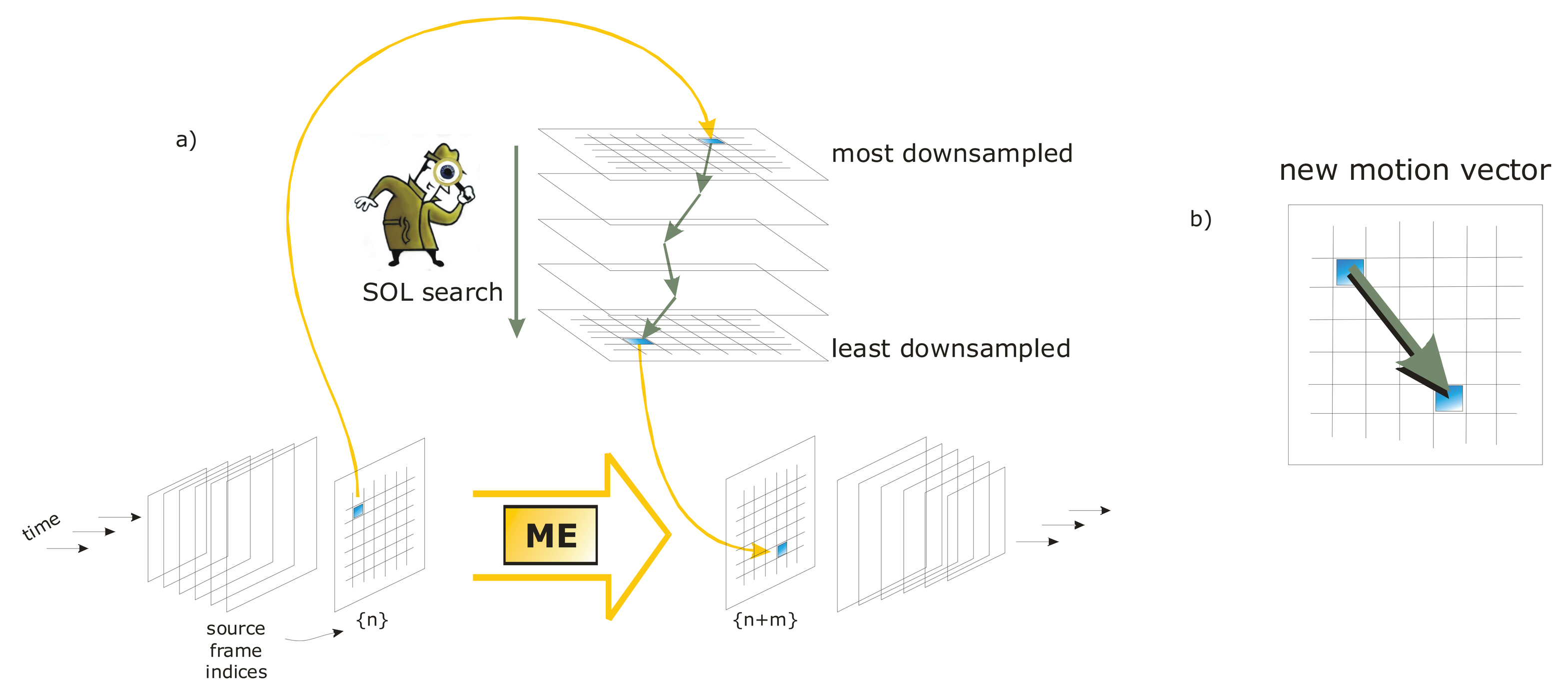

Figure 5. Motion Estimation is done based on Sum-Of-Luminance (SOL) values at each pixel and results in motion vectors for every pixel in frame {n} that map to positions in the {n+m} target frame. The blue pixel marks a single motion estimation process that involves a 5-level hierarchy pixel-wise search for the closest SOL value in the next frame. The hierarchy is comprised of 5 frames which are each down-sampled copies of the {n+m} target source frame. The levels represent varying levels of SOL down-sampling where the top level is the most down-sampled, and the bottom level is not down-sampled (it is an exact copy of the {n+m} frame). The SOL search proceeds down through the hierarchy until the closest SOL match is found, at which point the start and end positions of the motion vector are known. At this point, ME can proceed to the next pixel in frame {n}.

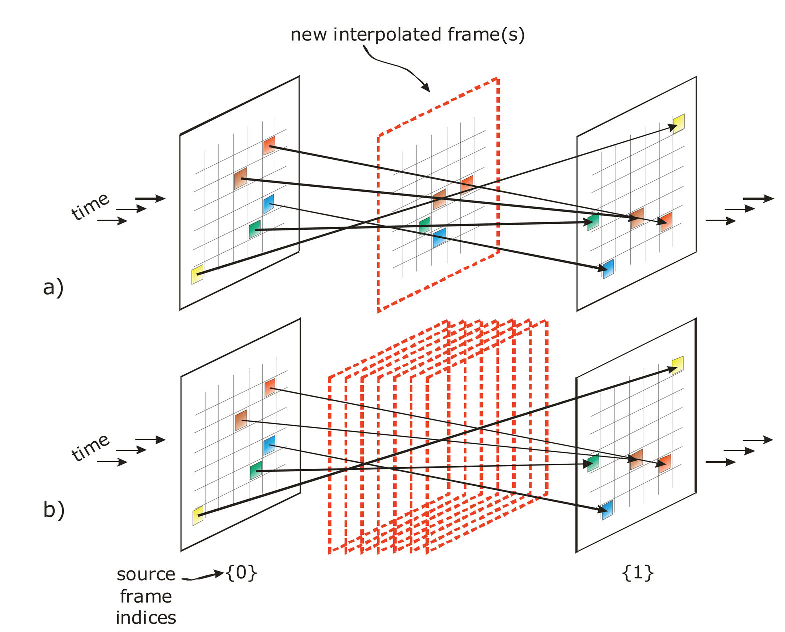

Figure 6. Pixel-wise MCI producing a) one, and b) many motion interpolated frames. The interpolation step uses motion information to guess where pixels (SOL value positions) will migrate over time across the "sweet-spot." Any number of interpolated frames can be produced because the motion vectors can pass through an infinite number of desired new frames within the sweet spot.

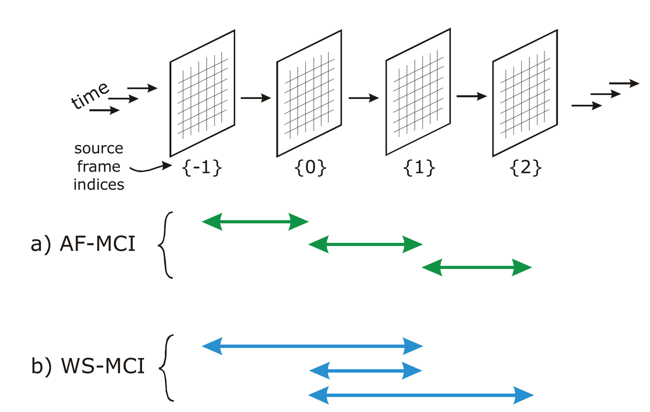

Figure 7. Comparison of Adjacent-Frame MCI (AF-MCI) and Wide-Span MCI (WS-MCI). The main difference is at the motion estimation step. AF only finds one set of motion vectors across the sweet-spot (a); whereas, WS uses three (b). Thereby, WS-MCI increases the exploitation of pertinent SOL-migration information across the time gap that is most motion-information rich (i.e. sweet spot).

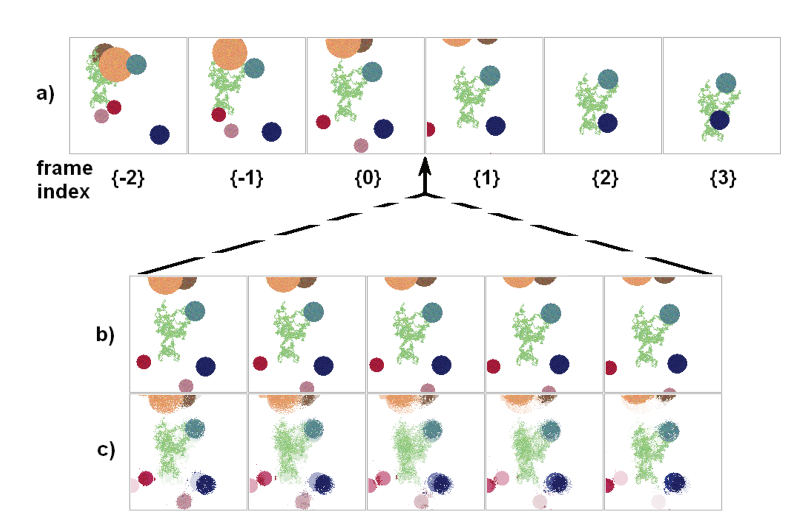

Figure 8. The general wide-span motion compensated interpolation (WS-MCI) assay uses an artificially generated a) source video sequence of 6 frames. These are used to produce both the b) 5 interpolated control frames and c) 5 experimental frames. This experimental run uses colour interpolation, and the source frames contain 7 random objects (inkblots or circles).

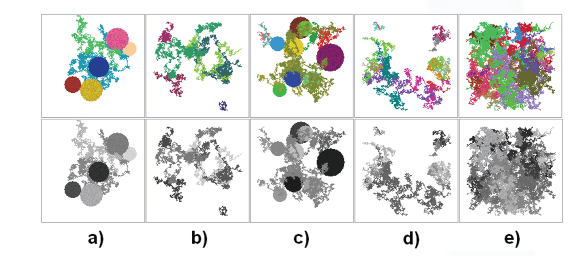

Figure 9. Representative source video frames (and their luminance-only frames below) showing a progression of source scene complexity. Each source frame is comprised of randomly positioned, textured, and scaled objects with random motion vectors. All these values are saved so controlled experimentation can be done. Frame a) includes 7 inkblots/circles, b) 7 inkblots, c) 14 inkblots/circles, d) 14 inkblots, and e) contains 21 inkblots. Recall that ME is performed on a hierarchy of these type of grey-scale frames.

Table 1. Statistical data comparing WS-MCI to AF-MCI over a progression of source video scene complexity. To rigorously analyze an MCI algorithm, an artificial video sequence was used to provide a reproducible and controlled sequence of data that could be directly compared to experimental results. Experimental frames were created for each test run and were directly compared to control frames in terms of the sum of their pixel colour differences for all interpolated frames. This allowed us to quantitatively assess the quality of a set of experimentally derived interpolated video frames. We have called this metric the *EC-colour-error* (experimental-control-colour-error). *WS-MCI % improvement* represents WS performance compared to AF-MCI.

Scene complexity	*# runs*	*AF-MCI av. EC-colour-err*	*WS-MCI av. EC-colour-err*	*Pearson correlation*	*2 tailed P-value*	*WS-MCI % improvmt*
1 circle	40	1.86	1.58	0.9850	9.17E-12	15.20%
7 circles	40	10.08	9.27	0.9812	6.63E-15	8.09%
7 objects	40	18.53	17.57	0.9982	3.65E-17	5.20%
7 inkblots	40	23.28	22.11	0.9983	1.13E-19	5.02%
14 circles	40	19.47	18.78	0.9768	3.11E-08	3.54%
14 objects	80	26.24	25.34	0.9949	5.07E-11	3.43%
14 inkblots	40	34.40	33.51	0.9958	2.57E-02	2.61%
21 inkblots	40	44.93	44.46	0.9865	3.00E-03	1.05%

Experimentation and statistical analysis shows a reproducible and significant advantage of WS-MCI over AF-MCI. For this reason, WS-MCI should be considered as part of any new motion compensation algorithm to produce better motion information and subsequently higher quality interpolated frames.

Future Work

Ultimately, we would like to apply WS-MCI to High-Definition TV (HDTV) signals. These signals present many challenges in terms of motion compensated interpolation: millions of pixels per frame, high colour precision, and encoding (ie. MPEG-2) for transmission. In addition to these technical issues, HDTV has a wider aspect ratio than typical NTSC TV. This means that more motion per frame is expected due to the fact that most video motion is horizontal.

Other future studies will investigate how well WS-MCI deals with non-linear motion, colour-guided ME, and backward motion vectors.

References

[1] B. Chupeau and E. Francois, "Region-based Motion Estimation for Content-based Video Coding and Indexing," *SPIE Visual Communications and Image Processing*, vol.4067, pp.884-893, 2000.

[2] B. Chupeau and P. Salmon, "Motion Compensating Interpolation for Improved Slow Motion," *Signal Processing of HDTV*, vol.4, pp.717-724, 1993.

[3] M. Tekalp, *Digital Video Processing*, Prentice Hall Inc., New Jersey, USA, 1995.

[4] W. Zheng, Y. Kanatsugu, S. Itoh, and Y. Tanaka, "Analysis of Space-dependent Characteristics of Motion-compensated Frame Differences," *International Conference on Image Processing*, vol.3, pp.158-164, 2000.

* submitted to the 2006 IEEE International Conference on Multimedia and Expo (IEEE ICME 2006), July 9-12 2006, Toronto, Canada.

† This work is funded by the Social Sciences and Humanities Research Council (SSHRC) of Canada through a Research Development Initiative grant under project number: 31-639484.