Loss concealment based on video inpainting for robust video communication

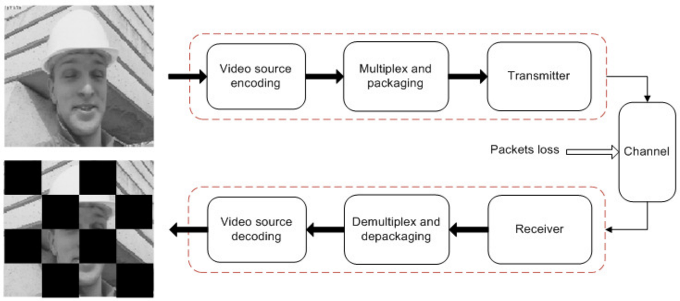
M. Ebdelli, O. Le Meur and C. Guillemot

August 30, 2012

Problematic and context

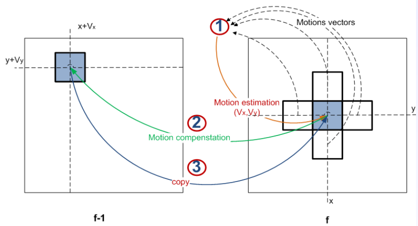
Error concealment :

Reconstruct lost blocks of received videos using the temporal and spatial correlation of video sequences.



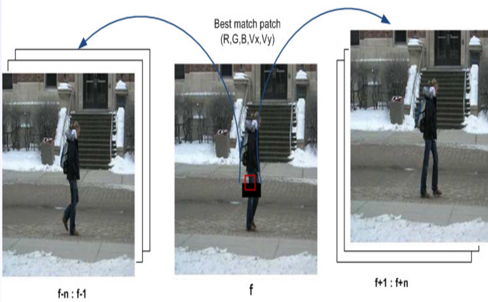
Two state-of-the-art approaches :

Motion-compensated temporal interpolation



1. Estimation of MV of lost block if it is not available.
2. Motion compensation.
3. Copy the motion compensated block.

Video inpainting



Fill-in missing blocks using patches similarities in the video.

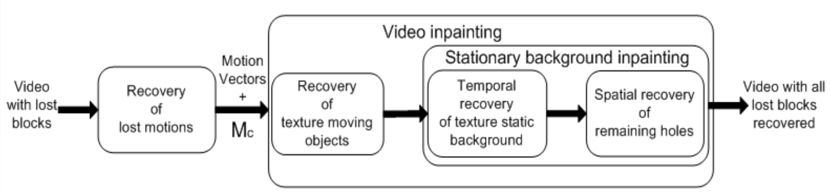
Problematic and context

1. Motion-compensated temporal interpolation :
   - ✔ Efficient for static part of images.
   - ✘ Error propagation : errors in motion estimation $\Rightarrow$ severe artefacts.
2. Examplar-based video inpainting :
   - ✔ Fill-in missing regions using motion and texture similarities.
   - ✘ motion vectors are necessary for better recovery.

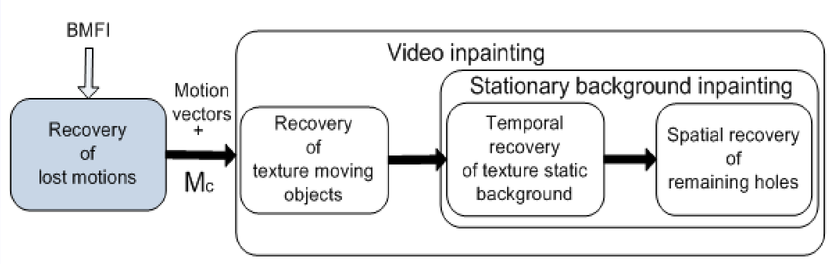
$\longrightarrow$ Proposed approach : combining temporal interpolation and video inpainting .

Overview of the proposed approach

- Consider the enhanced version of video inpainting algorithm [Patwardhan et al. (2005)] , proposed in [Ebdelli et al. (2012)].
- Motivations :
  1. Preprocessing step : estimation of motion vectors and moving pixels (M_c).
  2. Motion-compensated window to limit search space for best-matching patches.
  3. New frames priority for motion inpainting.



Bilinear Motion Field Interpolation (BMFI)

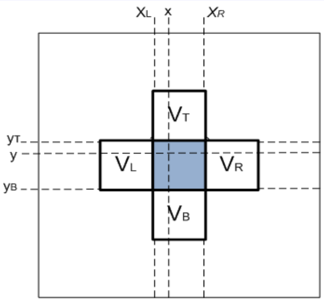


Bilinear Motion Field Interpolation (BMFI)

- Estimates the motion vector of each pixel p(x, y) in the lost block with bilinear interpolation of MV of neighboring blocks :

$$V(x,y) = \frac{1}{2}((1-x_n)V_L + x_n V_R + (1-y_n)V_T + y_n V_B)$$

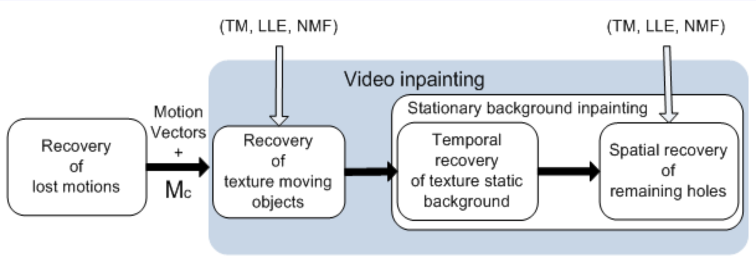
$$x_n = \frac{x-x_L}{x_R-x_L} \text{ and } y_n = \frac{y-y_T}{y_B-y_T}.$$



→ More accurate estimation for different types of motion.

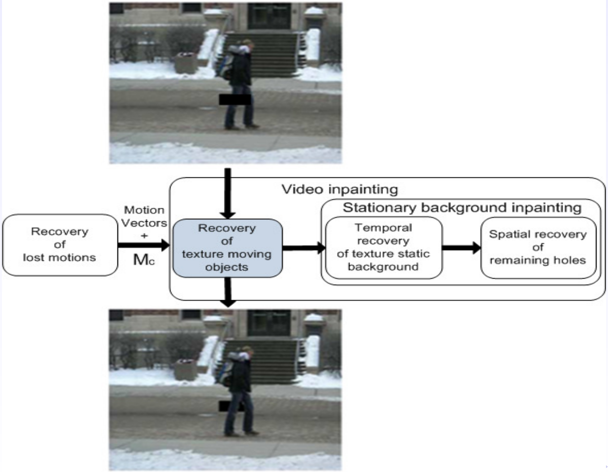
Video inpainting algorithm

Three step-algorithm :



Inpainting the moving foreground object

Only moving pixels of lost blocks are inpainted.



Inpainting the moving foreground object

Init : $C(p) = \begin{cases} 0 & \text{if p damaged.} \\ 1 & \text{otherwise.} \end{cases}$, $M_c(p) = \begin{cases} 1 & \text{if p is moving.} \\ 0 & \text{otherwise.} \end{cases}$

Repeat

Find the highest priority image **F** for motion filling

$$F = \arg\min(corrupted(M_c))$$

Repeat

Find the highest priority patch $\Psi_p \in$**F** for motion filling

$$p = \arg\max(C(p)D(p))$$

$$C(p) = \frac{\sum_{i \in \Psi_p} C(i)}{||\Psi_p||}, \quad D(p) = \frac{|\nabla M_c^{\perp} \cdot n_p|}{255}$$

Find the best matching patch Ψ_q to $\Psi_p, (R, G, B, V_x, V_y)$

Fill-in missing moving pixels of Ψ_p with their colocated in Ψ_q

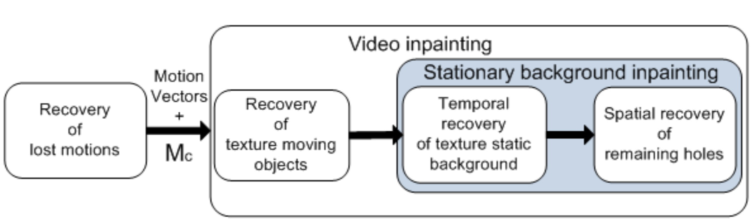
Untill no more patch need motion filling-in.

Untill no more image need motion filling-in.

Inpainting the stationary background

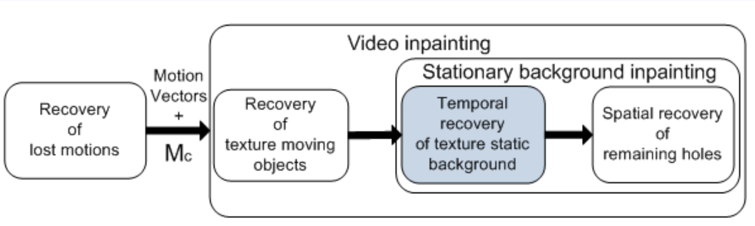
Inpaint damaged parts with stationary background in two step :

1. Inpaint damaged parts with stationary information of neighboring images.
2. Inpaint remaining holes with information from the same image : spatial inpainting.



Temporal inpainting of the stationary background

Inpaint damaged parts with stationary information of neighboring images.



Temporal inpainting of the stationary background

Init : $C(p) = \begin{cases} 0 & \text{if p is damaged.} \\ 1 & \text{otherwise.} \end{cases}$, $M(p) = \begin{cases} 0 & \text{if p is moving or damaged.} \\ 1 & \text{otherwise.} \end{cases}$

Repeat

Find the highest priority patch Ψ_p in the video for temporal filling.

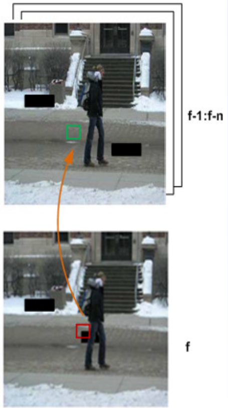
$$\{F, p\} = \arg\max(C(p).D(p))$$

$$C(p) = \frac{\sum_{i \in \Psi_p} C(i)}{\|\Psi_p\|} \, , \, D(p) = \frac{\sum_{i=F-n}^{F+n} M_i(p)}{2n+1}$$

Find the highest confident patch $\Psi_{p\,F_i}$ in neighboring frames of **F**.

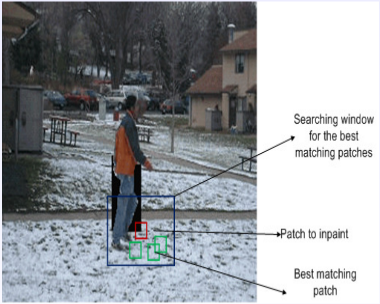
Copy non moving pixels of $\Psi_{p\,F_i}$ to Ψ_p.

Untill no temporal info can be copied.



Spatial inpainting

- No temporal information to copy from neighboring images → the hole has the same form in all images.
- Inpaint remain hole of each image using enhanced version of image inpainting algorithm [Criminisi et al. (2003)]:
  - search best matching patch of the patch to be inpainted using combinaition of (TM, LLE, NMF)[Ebdelli et al. (2012)].
- Each inpainted patch is copied to all images to have coherent result.



Performances

Original video	Corrupted video

- 20% of blocks of 16×16 are lost.
- Patch size for inpainting : 7×7.
- Searching window : 30×30.
- Resolution : 320×240.

Performances

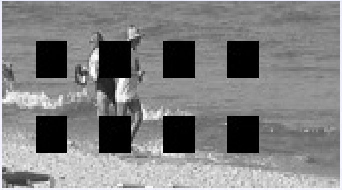
Spatial inpainting	Motion compensation MV estimated with BMFI	Proposed method
PSNR = 23.45 dB	PSNR = 31.47 dB	PSNR = 34.17 dB

Performances

Corrupted video	Spatial inpainting PSNR = 23.84 dB
Motion compensation PSNR = 30.74 dB	Proposed method PSNR = 33.39 dB

- 20% of blocks of 16×16 are lost.
- Patch size for inpainting : 7×7.
- Searching window : 30×30.
- Resolution : 320×240.

Performances

Original image	Corrupted image
	
Proposed method	Spatio-temporal selective extrapolation method [Meisinger and Kaupp (2007)]
	

Performances

PSNR (dB) values of recovering results using differents concealment methods :

Videos	Percentage of lost blocks	Spatial inpainting	Motion compensation using motions estimated with BMFI	Proposed method
1.1	20	23.45	31.47	34.17
1.2	50	19.37	27.27	29.27
2.1	20	23.84	30.74	33.39
2.2	50	19.90	27.13	28.84

$\rightarrow$ Average gain between concealed results using the proposed method and motion compensated interpolation is : 2.26 dB.

Conclusion and future work

1. Main contributions:
   - Combining motion interpolation and examplar-based video inpainting :
     $\rightarrow$ using texture similarity and motion information for better results
     $\rightarrow$ limits error propagation caused by the uncertainties on the estimated motion information.
   - Reduce complexity using a motion compensated window.
2. Future work:
   - Some artefacts still appear due to errors in estimating motion vectors of lost blocks $\rightarrow$ need of more robust method for motion vectors estimation of corrupted blocks.
   - Extend work to moving camera videos.

Thank you for your attention

Any questions?

*Bibliography

Criminisi, A., Perez, P., and Toyama, K. (Dec. 2003). Object removal by examplar-based image inpainting. In *In Proc. of International Conference on Computer Vision and pattern Recognition*, pages 721–728. CVPR.

Ebdelli, M., Guillemot, C., and Meur, O. L. (2012). Examplar-based video inpainting with motion-compensated neighbor embedding. In *Proc. of the IEEE Intl. Conf. on Image Processing, ICIP*.

Meisinger, K. and Kaupp, A. (Sept. 2007). Spatio-temporal selective extrapolation for 3-d signals and its applications in video communication. In *IEEE Trans. On Image Processing*, volume 16(9), pages 2348–2360.

Patwardhan, K. A., Sapiro, G., and Bertalmio, M. (2005). Video inpainting of occluding and occluded objects. In *Proc. of the IEEE Intl. Conf. on Image Processing, ICIP*, volume 2, pages 69–72.