Adapting Hierarchical ALS Algorithms for Temporal Psychovisual Modulation

Zhongpai Gao, Guangtao Zhai, Xiao Gu
Institute of Image Communication and Information Processing
Shanghai Jiao Tong University, Shanghai, China
Email: {gaozhongpai, zhaiguangtao, gugu97}@sjtu.edu.cn

Jiantao Zhou
Faculty of Science and Technology
University of Macau, Macau, China
Email: jtzhou@umac.mo

***Abstract*—Temporal psychovisual modulation (TPVM) is a new information display technology, which aims to generate multiple visual percepts for different viewers on a single display simultaneously. In TPVM system, the viewers with different active liquid crystal (LC) glasses (i.e., different modulation weights) which are synchronized with the display can see different images (called *personal views*). TPVM can be implemented by nonnegative matrix factorization (NMF) with three additional constrains: 1) the values of images and modulation weights should have upper bound; 2) a special view (called *shared view*) without using viewing devices should be considered (i.e., the sum of all basis images should be a meaningful image); 3) the sparsity of modulation weights should be considered because of the material property of LC. In this paper, we solve the constrained NMF problem by the modified hierarchical alternating least squares (HALS) algorithms. Through experiments, we analyse the influence of different parameters of TPVM to provide a guideline for parameter selection. This paper will provide an algorithmic guidance for the applications of TPVM.**

I. Introduction

Temporal psychovisual modulation (TPVM) [1], [2] was proposed as a new information display technology using the interplay of signal processing, optoelectronics and psychophysics. TPVM aims to generate multiple visual percepts for different viewers on a common exhibition medium concurrently. The rationale behind TPVM is that, on one hand, the human visual system (HVS) cannot resolve temporally rapidly changing optical signals beyond flicker fusion frequency (about 60 Hz for most of people). On the other hand, nowadays, modern display can work with 120 Hz or even higher refresh rate, so it is possible for a single display to provide different contents to different observers simultaneously. TPVM can be implemented by a combination of a high-speed display and display synchronized active liquid crystal (LC) glasses. The LC glasses are light blockage devices that can control how much of incoming light to pass through and enter the viewers eyes. A TPVM based display device broadcasts a set of images called atom frames at a speed higher than the flicker fusion frequency. The atom frames are then weighted by LC shutter based viewing devices that are synchronized with the display before entering the human visual system and integrating into some desired visual signals. Therefore, through different viewing devices, people can see different contents on the same display.

TPVM can be mathematically modeled as follows. Let f_d be the flicker fusion frequency and $f_c = Mf_d$ be the refresh rate of a digital display. So the display can emit $M = fc/fd$ atom frames, denoted as $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M$ ($\mathbf{x}_i \in \mathbb{R}^N$ for $i = 1, 2, \ldots, M$ and $N = m \times n$ is the resolution of each frame). The light fields of these M atom frames pass through and get amplitude-modulated by an active LC glasses. The M modulated atom frames are temporally fused by HVS and perceived as an image. Suppose that $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K$ ($\mathbf{y}_i \in \mathbb{R}^N$ for $i = 1, 2, \ldots, K$) are K target images formed by HVS. Then there should be K kinds of LC glasses with modulation weights $\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_K$ ($\mathbf{w}_i \in \mathbb{R}^M$ for $i = 1, 2, \ldots, K$). That is, the K target images $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_K$ perceived by viewers can be expressed as different linear combinations of the atom frames $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M$ (essentially basis functions) which means $\mathbf{Y} = \mathbf{XW}$, where $\mathbf{Y} \in \mathbb{R}^{N \times K}$, $\mathbf{X} \in \mathbb{R}^{N \times M}$ and $\mathbf{W} \in \mathbb{R}^{M \times K}$ contain all $\mathbf{y}_i$'s, $\mathbf{x}_i$'s and $\mathbf{w}_i$'s as their columns respectively. In this way, TPVM is a problem of signal decomposition $\mathbf{Y} = \mathbf{XW}$, as shown in Fig. 1.

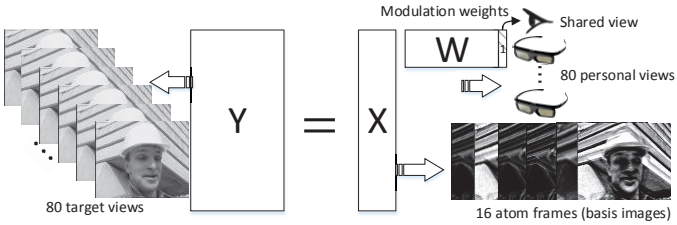

Fig. 1. Image formation by temporal psychovisual modulation (TPVM).

Since the light energy cannot be negative and active LC glasses cannot implement negative weights, the signal decomposition $\mathbf{Y} = \mathbf{XW}$ has to be a nonnegative matrix factorization (NMF). Moreover the modulation weight and the normalized gray scale value of image have the upper bound 1. So the TPVM system solves the following NMF with additional upper bound constraints optimization problem:

$$\min_{0 \le \mathbf{X}, \mathbf{W} \le 1} \frac{1}{2} \|s\mathbf{Y} - \mathbf{XW}\|_F^2, \tag{1}$$

where $\|\cdot\|_F$ is the Frobenius norm used to measure the distance between the target images and the reconstructed ones. The scaling factor $s \in [1, M]$ is to ensure adequate intensity of the images to be formed by TPVM. If M is sufficiently large, i.e., high refresh frequency f_d, the reconstruction errors can be small enough. For example, when $M = K$, the TPVM problem degenerates into the problem of temporal multiplexing ($\mathbf{X} = \mathbf{Y}, s = 1$, $\mathbf{W} = \mathbf{I}$, $\mathbf{I} \in \mathbb{R}^{M \times M}$ is an identity matrix), which has no reconstruction errors. However, when we achieve $K > M$ and $s > 1$, the TPVM display supports more views with brighter frames.

Visual exhibitions $\mathbf{y}_i$'s produced by the TPVM display system require the use of synchronized active shutter glasses.

978-1-4799-8391-9/15/$31.00 ©2015 IEEE

However, another image will also be generated to the viewers who do not use light modulation devices. This view formed by fusing all unattenuated atom frames, i.e., $s_0\mathbf{y}_0 = \mathbf{x}_1 + \mathbf{x}_2 + \ldots + \mathbf{x}_M$, is the default image seen by all viewers without using personalized viewing devices, $s_0 \in [s, M]$ is also to ensure adequate intensity of the images formed by TPVM. In many multiuser applications, the *shared view* $\mathbf{y}_0$ should be semantically meaningful to be deceptive or visual pleasing to the bystanders. Once considered the *shared view* $\mathbf{y}_0$, the objective function (1) can be expanded as

$$\min_{0\le\mathbf{X},\mathbf{W}\le 1} \frac{1}{2}\|s\mathbf{Y} - \mathbf{XW}\|_F^2 + \frac{\lambda_{sh}}{2}\|s_0\mathbf{y_0} - \mathbf{X1}\|_F^2, \quad (2)$$

where $\mathbf{1}$ stands for a column vector of all 1s and λ_{sh} is a nonnegative regularization coefficient controlling quality tradeoff between the *shared view* $\mathbf{y}_0$ and *personal views* $\mathbf{y}_i$, $1 \le i \le K$.

Considering the material property and operating mode of L-C, the speed of the off-the-shelf LC glasses cannot be too high, meaning in mathematics a fewer number of large elements of the modulation weights. This can be achieved by adding sparsity constraint to the objective function (2) by making $\|\mathbf{W}\|_1$ as small as possible where $\|\mathbf{W}\|_1 = \sum_{mk} \mathbf{W}_{mk}$, expressly,

$$\min_{0\le\mathbf{X},\mathbf{W}\le 1} \frac{1}{2}\|s\mathbf{Y} - \mathbf{XW}\|_F^2 + \frac{\lambda_{sh}}{2}\|s_0\mathbf{y}_0 - \mathbf{X1}\|_F^2 + \lambda_{sp}\|\mathbf{W}\|_1, \quad (3)$$

where λ_{sp} is a nonnegative regularization coefficient controlling sparsity of the matrix $\mathbf{W}$ [2].

The above constrained NMF problem can be solved via modifying existing NMF algorithms. The NMF problem

$$\min_{\mathbf{X},\mathbf{W}\ge 0} \frac{1}{2}\|\mathbf{Y} - \mathbf{XW}\|_F^2, \quad (4)$$

is a nonconvex optimization problem with respect to $\mathbf{W}$ and $\mathbf{X}$, which can just lead to a local minimum. Hence, most of known algorithms for the NMF model are based on alternating least squares (ALS) minimization of the Squared Euclidean distance. That is, NMF reduces to an efficiently solvable convex nonnegative least squares (NNLS) problem when one of the factors $\mathbf{W}$ or $\mathbf{X}$ keeps fixed. The general framework of ALS can be briefly described as follow:

1) Initial matrix $\mathbf{W}$ with nonnegative elements.
2) Repeat solving the following NNLS problems

$$\min_{\mathbf{X}\ge 0} \frac{1}{2}\|\mathbf{Y} - \mathbf{XW}\|_F^2, \quad (5)$$

$$\min_{\mathbf{W}\ge 0} \frac{1}{2}\|\mathbf{Y}^T - \mathbf{W}^T\mathbf{X}^T\|_F^2, \quad (6)$$

until a convergence criterion is satisfied.

A class of algorithms called alternating nonnegative least squares(ANLS) [3] is based on this framework. It computes an optimal solution for the NNLS subproblem. However the computation of this active-set-like method is very costly. Some algorithms which only compute an approximate solution of NNLS subproblem are explored. These algorithms sometimes are rough but with lower computational costs. The multiplicative update (MU) algorithms proposed by Lee and Seung [4] is the first well known NMF algorithms. But they are slow and even do not guarantee to converge [5]. To speed up the convergence of MU, some projected gradient bound-constrained optimization methods have been proposed [5]. The truncated alternating least squares method (TALS) [6] just sets all negative elements resulting from the least squares computation to 0. But the TALS algorithm is easy to get stuck in local minima and difficult to convergent. The hierarchical alternating least squares (HALS) introduced by Cichocki *et al.* [7] can alleviate the problem of getting stuck in local minima by improved local leaning rules (i.e., columns of matrices are processed sequentially one by one). In this paper, we modify the HALS algorithms to solve the TPVM constrained NMF problem in (3) [8].

II. Hierarchical ALS Algorithms

In HALS, instead of minimizing one or two cost functions, the columns of $\mathbf{X}$ and $\mathbf{W}^T$ are updated sequentially. That is, we update a single column of $\mathbf{X}$ where fixing all the other variables. We define the residues

$$\mathbf{Y}^{(p)} = \mathbf{Y} - \sum_{j\ne p} \mathbf{X}_{:j}\mathbf{W}_{j:} = \mathbf{Y} - \mathbf{XW} + \mathbf{X}_{:p}\mathbf{W}_{p:}, \quad (7)$$

where $\mathbf{Z}_{:p}$ and $\mathbf{Z}_{p:}$ denote the pth column and pth row of matrix $\mathbf{Z}$ respectively. So the problem in (5) reduces to

$$\min_{\mathbf{X}_{:p}\ge 0} \frac{1}{2}\|\mathbf{Y} - \mathbf{XW}\|_F^2 = \frac{1}{2}\|\mathbf{Y}^{(p)} - \mathbf{X}_{:p}\mathbf{W}_{p:}\|_F^2 \quad (8)$$

for $p = 1, 2, \ldots, M$. The gradients of the local cost functions $D_F^{(p)}(\mathbf{Y}^{(p)}\|\mathbf{X}_{:p}\mathbf{W}_{p:})$ in (8) can be expressed by

$$\frac{\partial D_F^{(p)}(\mathbf{Y}^{(p)}\|\mathbf{X}_{:p}\mathbf{W}_{p:})}{\partial \mathbf{X}_{:p}} = \mathbf{X}_{:p}\mathbf{W}_{p:}\mathbf{W}_{p:}^T - \mathbf{Y}^{(p)}\mathbf{W}_{p:}^T. \quad (9)$$

Then we can obtain the sequential learning rules by equating the gradients to zero and setting the negative elements to zero:

$$\begin{aligned}\mathbf{X}_{:p} &\leftarrow [\mathbf{Y}^{(p)}\mathbf{W}_{p:}^T(\mathbf{W}_{p:}\mathbf{W}_{p:}^T)^{-1}]_+ \quad (10)\\ &= [(\mathbf{Y} - \mathbf{XW} + \mathbf{X}_{:p}\mathbf{W}_{p:})\mathbf{W}_{p:}^T(\mathbf{W}_{p:}\mathbf{W}_{p:}^T)^{-1}]_+\\ &= [([\mathbf{YW^T}]_{:p} - \mathbf{X}[\mathbf{WW}^T]_{:p} + \mathbf{X}_{:p}\mathbf{W}_{p:}\mathbf{W}_{p:}^T)(\mathbf{W}_{p:}\mathbf{W}_{p:}^T)^{-1}]_+\end{aligned}$$

where $[\mathbf{Z}]_+ = \max\{\mathbf{0}, \mathbf{Z}\}$ is the component-wise maximum. We denote $\mathbf{A} = \mathbf{YW}^T$ and $\mathbf{B} = \mathbf{WW}^T$. Then the update rule in (10) can be simplified as

$$\mathbf{X}_{:p} \leftarrow [\frac{\mathbf{A}_{:p} - \mathbf{XB}_{:p} + \mathbf{X}_{:p}\mathbf{B}_{pp}}{\mathbf{B}_{pp}}]_+ = [\mathbf{X}_{:p} + \frac{\mathbf{A}_{:p} - \mathbf{XB}_{:p}}{\mathbf{B}_{pp}}]_+. \quad (11)$$

Analogous to equation (11), the learning rule for a single row $\mathbf{W}_{p:}$ of $\mathbf{W}$ in (6) is given by

$$\mathbf{W}_{p:} \leftarrow [\mathbf{W}_{p:} + \frac{\mathbf{C}_{p:} - \mathbf{D}_{p:}\mathbf{W}}{\mathbf{D}_{pp}}]_+, \quad (12)$$

where $\mathbf{C} = \mathbf{X}^T\mathbf{Y}$ and $\mathbf{D} = \mathbf{X}^T\mathbf{X}$.

By evaluating the computational cost needed at each iteration (11) or (12), Gillis and Glineur [9] proposed a simple way that significantly accelerates these schemes. Suppose $NK > MN + MK$ (i.e., the number of entries in $\mathbf{X}$ and $\mathbf{W}$ is smaller than that in $\mathbf{Y}$), the computation of $\mathbf{A}$ or $\mathbf{C}$ is the most expensive among the learning rules in (11) or (12). Therefore, the time-consuming step should be performed sparsely to reduce the computational costs. This can be achieved by updating (11) or (12) several times before the next update (12) or (11). The numbers of inner iterations of (11) and (12) can be determined by the flop counts and the supplementary stopping criterion described in [9].

III. Modified HALS for TPVM

The above HALS algorithms can be extended to the TPVM problem in (3) by imposing additional constrains (upper bound, *shared view* and sparsity constrains). The constrained NMF can be reduced to the following NNLS problems

$$\min_{0\le\mathbf{X}\le 1}\frac{1}{2}\|s\mathbf{Y}-\mathbf{X}\mathbf{W}\|_F^2+\frac{\lambda_{sh}}{2}\|s_0\mathbf{y}_0-\mathbf{X}\mathbf{1}\|_F^2, \tag{13}$$

$$=\frac{1}{2}\|\tilde{\mathbf{Y}}-\mathbf{X}\tilde{\mathbf{W}}\|_F^2 \tag{14}$$

$$\min_{0\le\mathbf{W}\le 1}\frac{1}{2}\|s\mathbf{Y}-\mathbf{X}\mathbf{W}\|_F^2+\lambda_{sp}\|\mathbf{W}\|_1, \tag{15}$$

where $\tilde{\mathbf{Y}}=[\sqrt{\lambda_{sh}}s_0\mathbf{y}_0|s\mathbf{Y}]$ and $\tilde{\mathbf{W}}=[\sqrt{\lambda_{sh}}\mathbf{1}_{M\times 1}|\mathbf{W}]$. Analogous to the HALS algorithms in section II, The gradients of the local cost functions can be expressed by

$$\frac{\partial D_{\mathbf{X}}^{(p)}}{\partial\mathbf{X}_{:p}}=\mathbf{X}_{:p}\tilde{\mathbf{W}}_{p:}\tilde{\mathbf{W}}_{p:}^T-\tilde{\mathbf{Y}}^{(p)}\tilde{\mathbf{W}}_{p:}^T, \tag{16}$$

$$\frac{\partial D_{\mathbf{W}}^{(p)}}{\partial\mathbf{W}_{p:}}=\mathbf{X}_{:p}^T\mathbf{X}_{:p}\mathbf{W}_{p:}-s\mathbf{X}_{:p}^T\mathbf{Y}^{(p)T}+\lambda_{sp}\mathbf{1}_{K\times 1}^T, \tag{17}$$

where

$$\tilde{\mathbf{Y}}^{(p)}=\tilde{\mathbf{Y}}-\mathbf{X}\tilde{\mathbf{W}}+\mathbf{X}_{:p}\tilde{\mathbf{W}}_{p:}. \tag{18}$$

Then the updating rules can be expressed as

$$\mathbf{X}_{:p}\leftarrow[\mathbf{X}_{:p}+\frac{\tilde{\mathbf{A}}_{:p}-\mathbf{X}\tilde{\mathbf{B}}_{:p}}{\tilde{\mathbf{B}}_{pp}}]_{[0,1]}, \tag{19}$$

$$\mathbf{W}_{p:}\leftarrow[\frac{s\mathbf{C}_{p:}-\mathbf{D}_{p:}\mathbf{W}+\mathbf{D}_{pp}\mathbf{W}_{p:}-\lambda_{sp}\mathbf{1}_{K\times 1}^T}{\mathbf{D}_{pp}}]_{[0.1]}, \tag{20}$$

where $[\mathbf{Z}]_{[0,1]}=\min(\max(\mathbf{Z},\mathbf{0}),\mathbf{1})$, $\tilde{\mathbf{A}}=\tilde{\mathbf{Y}}\tilde{\mathbf{W}}^T$ and $\tilde{\mathbf{B}}=\tilde{\mathbf{W}}\tilde{\mathbf{W}}^T$. The pseudocode of the modified HALS for TPVM is displayed in Algorithm 1.

Algorithm 1 Modified HALS for TPVM

Require: $\mathbf{Y}\in\mathbb{R}^{N\times K}$

Initialize nonnegative matrices $\mathbf{W}\in\mathbb{R}^{M\times K},\mathbf{X}\in\mathbb{R}^{N\times M}$
$\tilde{\mathbf{Y}}=[\sqrt{\lambda_{sh}}s_0\mathbf{y}_0|s\mathbf{Y}]$; $\tilde{\mathbf{W}}=[\sqrt{\lambda_{sh}}\mathbf{1}_{M\times 1}|\mathbf{W}]$;
$\tilde{\mathbf{A}}=\tilde{\mathbf{Y}}\tilde{\mathbf{W}}^T$; $\tilde{\mathbf{B}}=\tilde{\mathbf{W}}\tilde{\mathbf{W}}^T$;
repeat
  % Update $\mathbf{X}$
  $\tilde{\mathbf{W}}\leftarrow[\sqrt{\lambda_{sh}}\mathbf{1}_{M\times 1}|\mathbf{W}]$;
  $\tilde{\mathbf{A}}\leftarrow\tilde{\mathbf{Y}}\tilde{\mathbf{W}}^T$; $\tilde{\mathbf{B}}\leftarrow\tilde{\mathbf{W}}\tilde{\mathbf{W}}^T$;
  repeat
    for $i=1$ to M **do**
      $\mathbf{X}_{:p}\leftarrow\mathbf{X}_{:p}+\frac{\tilde{\mathbf{A}}_{:p}-\mathbf{X}\tilde{\mathbf{B}}_{:p}}{\tilde{\mathbf{B}}_{pp}}$;
      $\mathbf{X}_{:p}\leftarrow\min(\max(\mathbf{X}_{:p},\mathbf{0}),\mathbf{1})$;
  until the number of inner iterations is reached
  % Update $\mathbf{W}$
  $\mathbf{C}\leftarrow\mathbf{X}^T\mathbf{Y}$; $\mathbf{D}\leftarrow\mathbf{X}^T\mathbf{X}$;
  repeat
    for $i=1$ to M **do**
      $\mathbf{W}_{p:}\leftarrow\frac{s\mathbf{C}_{p:}-\mathbf{D}_{p:}\mathbf{W}+\mathbf{D}_{pp}\mathbf{W}_{p:}-\lambda_{sp}\mathbf{1}_{K\times 1}^T}{\mathbf{D}_{pp}}$;
      $\mathbf{W}_{p:}\leftarrow\min(\max(\mathbf{W}_{p:},\mathbf{0}),\mathbf{1})$;
  until the number of inner iterations is reached
until convergence criterion is reached

TABLE I. R and R_0 for the HALS algorithms using different values of K with $s=1,s_0=3,\lambda_{sh}=2,\lambda_{sp}=2$.

$K=$	M	$2M$	$3M$	$4M$	$5M$	$6M$	$7M$	$8M$
$R(\%)$	**1.68**	3.30	4.38	5.31	6.02	6.81	7.00	7.11
$R_0(\%)$	**0.28**	0.81	1.25	1.56	1.73	2.29	2.41	2.83

IV. Experiments and Results

In this section, we simulate the potential performance of TPVM using the modified HALS algorithms introduced in section III on a 3.4 GHz Intel(R) Core(TM) i7-2600K CPU Windows 64-bit PC with 16 GB memory. The sequences of YUV video 'foreman'[1] with resolution 352×288 are the test data for simulation. The TPVM system in (3) requires 5 parameters: the number of target views K (i.e., the first K frames of 'foreman'), scaling factors s and s_0 and regularization coefficients λ_{sh} and λ_{sp}. The number of atom frames M is determined by the refresh rate f_d of the display and suppose $M=16$ in our experiments. Relative object value (ROV) is chosen to be the performance index. ROV of the *personal views* $\mathbf{Y}$ is defined as

$$R(s\mathbf{Y},\mathbf{X}\mathbf{W})=\|s\mathbf{Y}-\mathbf{X}\mathbf{W}\|_F/\|s\mathbf{Y}\|_F. \tag{21}$$

And the ROV of the *shared view* $\mathbf{y}_0$ is

$$R_0(s_0\mathbf{y}_0,\mathbf{X}\mathbf{1})=\|s_0\mathbf{y}_0-\mathbf{X}\mathbf{1}\|_F/\|s_0\mathbf{y}_0\|_F. \tag{22}$$

ROV measures the similarity of reconstructed images to the target views and the smaller ROV is the better. For convenience and without losing generality, we set $\mathbf{y}_0=\mathbf{y}_1$ in our experiments. The influence of each parameter is discussed in the following paper through experiments.

Target views number K: for TPVM system, to provide sufficient target views, the larger K is the better. However, as K increases, the relative object errors R and R_0 also increase, as shown in TABLE I. There is a tradeoff between the number of target views and the similarity of reconstructed images to the target views. That is, under a certain fidelity of reconstructed images, we choose the largest value of K. In our experiments, we set $K=80=5M$.

Scaling factors s *and* s_0: generally, light intensity of the *shared view* $\mathbf{y}_0$ is larger than the *personal views* $\mathbf{y}_i(1<i\le K)$ because some light of the *personal views* is blocked by LC glasses. This can be expressed as $s_0>s$ mathematically. We fix s and let $R_0=(1.72\pm 0.2)\%$ which ensures the fidelity of the *shared view*. Then we set $s_0=ks(1\le k\le 8)$ as shown in TABLE II. We find that R is negatively correlated to s_0. That is, the greater brightness difference of the *shared view* and the *personal views*, the higher fidelity of the views. However, the brightness of the *shared view* and the *personal views* is expected to be close to each other in practice. So we set $s2=3S$ in our experiments. R and R_0 is positively correlated to s in general as shown in TABLE III. So we set $s=1$. The scaling factors can be larger properly to increase the light intensity of the views in practice.

Regularization coefficient λ_{sh}: from TABLE IV and Fig. 2, we can find that R and R_0 are positively correlative and negatively correlative to λ_{sh} respectively. So there is a quality tradeoff between the *shared view* and the *personal views*.

[1]YUV video 'foreman': http://trace.eas.asu.edu/yuv/index.html

TABLE II. R FOR THE HALS ALGORITHMS USING DIFFERENT VALUES OF s_0 WITH $R_0 = (1.72 \pm 0.2)\%$, s = 1, $\lambda_{sp} = 2$, $K = 5M$.

$s_0 =$	s	$2s$	$3s$	$4s$	$5s$	$6s$	$7s$	$8s$
λ_{sh}	600	9	2	0.78	0.43	0.25	0.22	**0.2**
$R(\%)$	17.75	6.71	6.01	5.78	5.67	5.61	5.60	**5.56**

TABLE III. R AND R_0 FOR THE HALS ALGORITHMS USING DIFFERENT VALUES OF s WITH $s_0 = 3s$, λ_{sh} = 2, $\lambda_{sp} = 2$, $K = 5M$.

$s =$	1	2	3	4	5	6	7	8
$R(\%)$	6.01	**5.97**	6.02	6.32	7.06	8.27	9.38	10.51
$R_0(\%)$	**1.73**	1.78	2.09	3.06	4.60	8.51	15. 89	23.29

TABLE IV. R AND R_0 FOR THE HALS ALGORITHMS USING DIFFERENT VALUES OF λ_{sh} WITH $s = 1$, $s_0 = 3s$, $\lambda_{sp} = 2$, $K = 5M$.

$\lambda_{sh} =$	0.0	0.5	1.0	1.5	2.0	2.5	3.0	3.5
$R(\%)$	**5.37**	5.84	5.91	5.94	5.99	6.00	6.03	6.06
$R_0(\%)$	19.74	4.31	3.02	2.33	1.82	1.65	1.34	**1.27**

TABLE V. R, R_0 AND THE SPARSITY OF W FOR THE HALS USING DIFFERENT VALUES OF λ_{sp} WITH $s = 1$, $s_0 = 3s$, $\lambda_{sh} = 2$, $K = 5M$.

$\lambda_{sp} =$	0	2	4	6	8	10	12	14
$R(\%)$	6.03	5.90	5.95	5.94	5.96	6.04	6.08	6.12
$R_0(\%)$	1.81	1.95	1.84	1.96	1.97	2.05	2.14	2.08
Sparsity(%)	20.1	22.0	25.9	26.6	28.4	20.6	35.8	**40.9**

However the influence of changing λ_{sh} to R is smaller than to R_0. That is, we can sacrifice the quality of the *personal views* a little to improve the quality of the *shared view* a lot. We set $\lambda_{sh} = 2$ in our experiments.

Regularization coefficient λ_{sp}: the sparsity of W can be defined as the proportion of the zero elements. The sparsity of W is positively correlative to λ_{sp} and the fidelity of the views decreases slightly when λ_{sp} grows as shown in TABLE V. So the sparsity can be set according to the need of actual situation.

The experiments results are shown in Fig. 3. More results and the code will be presented on the website: http://multimedia.sjtu.edu.cn/.

V. CONCLUSION

In this paper we formulate temporal psychovisual modulation (TPVM) by nonnegative matrix factorization (NMF) with additional constrains. And we adapt the hierarchical alternating least squares (HALS) algorithms to generate the atom frames and modulation weights for TPVM. Through experiments, we analyse the influence of five parameters of TPVM to provide a guideline for parameter selection. The five parameters include the number of target views K , scaling factor of the *personal views* light intensity s and the *shared view* light intensity s_0, the *shared view* regularization coefficient λ_{sh} and modulation weights sparsity regularization coefficient λ_{sp}. This paper will provide an algorithmic guidance for the applications of TPVM.

ACKNOWLEDGMENT

This work was supported in part by NSFC (61422112, 61331014, 61371146) and Macau Science and Technology Development Fund under Grant FDCT 046/2014/A1.

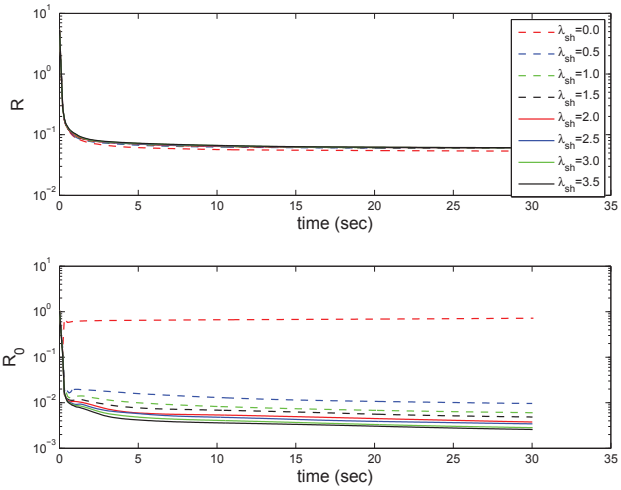

Fig. 2. Relative Objective Value R and R_0 versus time for different λ_{sh} with $s = 1, s_0 = 3s, \lambda_{sp} = 2, K = 5M$.

Fig. 3. Comparison of the target views and reconstructed images. The first row are the *shared view* and three *personal views* with distinctive contents in 'foreman' and the second row are the reconstructed images. We set $s = 1, s_0 = 3s, \lambda_{sh} = 2, \lambda_{sp} = 2, K = 5M$.

REFERENCES

[1] X. Wu and G. Zhai, "Temporal psychovisual modulation: A new paradigm of information display [Exploratory DSP]," *Signal Processing Magazine, IEEE*, vol. 30, no. 1, pp. 136–141, 2013.

[2] G. Zhai and X. Wu, "Multiuser collaborative viewport via temporal psychovisual modulation [Applications Corner]," *Signal Processing Magazine, IEEE*, vol. 31, no. 5, pp. 144–149, 2014.

[3] J. Kim and H. Park, "Fast nonnegative matrix factorization: An active-set-like method and comparisons," *SIAM Journal on Scientific Computing*, vol. 33, no. 6, pp. 3261–3281, 2011.

[4] D. D. Lee and H. S. Seung, "Learning the parts of objects by non-negative matrix factorization," *Nature*, vol. 401, no. 6755, pp. 788–791, 1999.

[5] E. F. Gonzalez and Y. Zhang, "Accelerating the Lee-Seung algorithm for non-negative matrix factorization," *Dept. Comput. & Appl. Math., Rice Univ., Houston, TX, Tech. Rep. TR-05-02*, 2005.

[6] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, "Algorithms and applications for approximate nonnegative matrix factorization," *Computational Statistics & Data Analysis*, vol. 52, no. 1, pp. 155–173, 2007.

[7] A. Cichocki and P. Anh-Huy, "Fast local algorithms for large scale nonnegative matrix and tensor factorizations," *IEICE transactions on fundamentals of electronics, communications and computer sciences*, vol. 92, no. 3, pp. 708–721, 2009.

[8] J. Feng, X. Huo, L. Song, X. Yang, and W. Zhang, "Evaluation of different algorithms of nonnegative matrix factorization in temporal psychovisual modulation," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 24, no. 4, pp. 553–565, April 2014.

[9] N. Gillis and F. Glineur, "Accelerated multiplicative updates and hierarchical ALS algorithms for nonnegative matrix factorization," *Neural computation*, vol. 24, no. 4, pp. 1085–1105, 2012.