

## Research Article

# A Perception-Driven Transcale Display Scheme for Space Image Sequences

Lingling Zi <sup>1</sup>, Xin Cong <sup>1</sup>, Yanfei Peng,<sup>1</sup> and Pei Yang<sup>2</sup>

<sup>1</sup>School of Electronic and Information Engineering, Liaoning Technical University, Huludao 125105, China

<sup>2</sup>College of Science, Liaoning Technical University, Fuxin 123000, China

Correspondence should be addressed to Xin Cong; [chongzi610@163.com](mailto:chongzi610@163.com)

Received 22 May 2018; Accepted 19 September 2018; Published 4 October 2018

Academic Editor: Deepu Rajan

Copyright © 2018 Lingling Zi et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of multimedia technology, the way of obtaining high-quality motion reproduction for space targets has attracted much attention in recent years. This paper proposes a Perception-driven Transcale Display Scheme, which significantly improves the awareness of multimedia processing. This new scheme contains two important modules, transcale description based on visual saliency and perception-driven display of space image sequences. The former concentrates on describing the transcale feature of space targets, including three algorithms, attention region computing, frame rate conversion, and image resolution resizing. On this basis, the latter focuses on high-quality display of space movements under different scales, including three algorithms, namely, target trajectory computing, space transcale display, and space movement display. Extensive quantitative and qualitative experimental evaluations demonstrate the effectiveness of the proposed scheme.

## 1. Introduction

With the rapid development of computer technology, spatial rendezvous and docking needs a large amount of image information between navigation information systems. As an important information form, space image sequences play an important role, such as docking control, docking mechanism design, and track interactive control, and they need to be transmitted between these navigation information systems to measure the relative position, relative velocity, and relative attitude of spacecraft [1–4]. For the spatial image sequence, transcale is its important feature. So, demanding multimedia processing methods are different from those of traditional image processing. At present, the study of space image sequences is still at the initial stage and further exploration is still on the way. Specifically, there are mainly two unresolved problems. The first one is how to achieve the task of high-quality monitoring space targets. For example, the continuous attention to space targets should be strengthened due to the lack of the details of space movements. The second one is how to improve the ability of processing images, such as smooth motion reproduction, accurate trajectory description, and different display effects; all of them should

be fine exhibited to enhance the awareness of multimedia processing. To solve the above problems, a transcale display method [5–7] emerged as a new direction for image processing and the challenges are mainly shown as follows.

(1) The change of the attention scale: The existing image processing methods only use the inherent characteristics of the image sequences and lack the attention perceived by observers, which needs to be converted from the description of the whole sequences to that of space targets, so as to better capture fine details of space image sequences.

(2) The change of the scale of the frame rate: The existing frame interpolation methods have the problem that the details of the interpolation frames are not clear during the process of changing the frame rates, which affects the smoothness of space image sequences. Thereby, new methods need to be studied to obtain the optimal interpolation frames.

(3) The change of the resolution scale: Because of the limited screen sizes of different devices, the same image sequence is frequently required to be displayed in the different sizes. However, the important contents are not changed the scales in the existing resizing methods, so new methods need to be proposed to improve the definition of important contents under the guarantee of the global resizing effects.

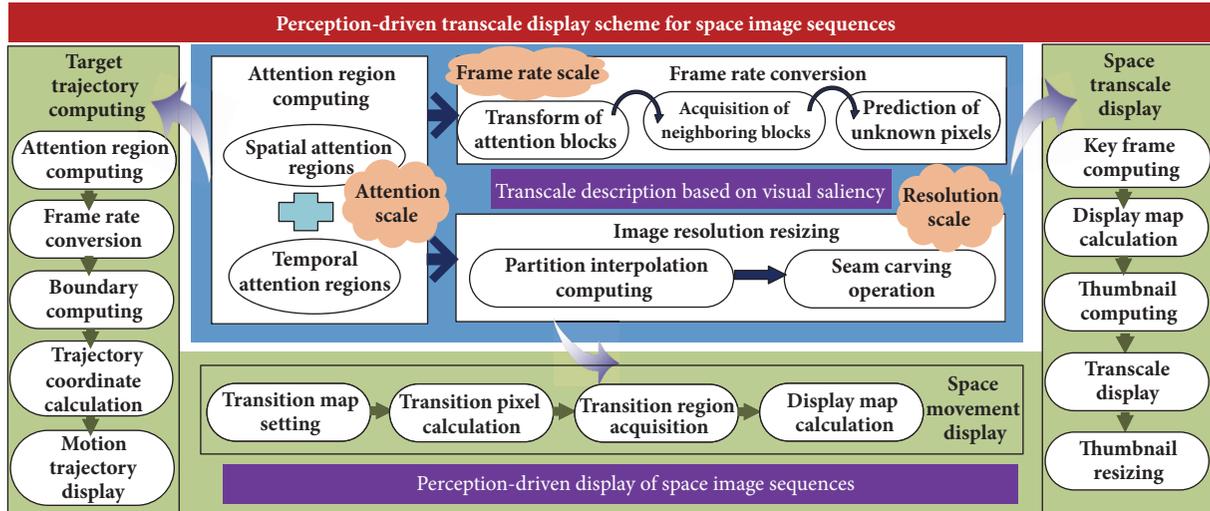


FIGURE 1: The framework of PTDS.

Facing the above challenges, we propose a Perception-driven Transcale Display Scheme (PTDS) to achieve high-quality space motion reproduction. The main contributions have three points. Firstly, we construct a transcale display framework by providing a different perspective, which significantly improves the awareness of multimedia processing. This framework contains two important modules, transcale description based on visual saliency and perception-driven display of space image sequences. Secondly, the module of transcale description is presented to solve the transcale problems, which is the core of PTDS. Finally, the module of perception-driven display is proposed to realize the trajectory description and movement display for space targets under the condition of changing scales. To sum, PTDS could serve for navigation information systems.

The rest of the paper is structured as follows. Section 2 discusses the framework of PTDS. Sections 3 and 4 describe in detail the formulation of PTDS, namely, transcale description and perception-driven display. Section 5 presents experimental work carried out to demonstrate the effectiveness of PTDS. Section 6 concludes the paper.

## 2. The Framework of PTDS

Figure 1 shows the framework of PTDS and it contains two important modules, transcale description and perception-driven display. The first one is to improve the definitions of the important contents perceived by observers under the condition of changing different scales, including the attention scale, the scale of frame rate, and the resolution scale. The way to determine the attention regions from the observers' viewpoint is naturally becoming a key issue. Recently, the visual saliency technique is more and more widely used in the multimedia field [8–11]. The reasons for using this technique in image processing contain the following three points. Firstly, it can provide the capability of choosing for image description, strengthen the description details of moving targets, so as to priority allocation to achieve the

desired image analysis and synthesis of computing resources. Secondly, it can accomplish auxiliary construction of the flexible description scheme by discriminating the saliency regions and non-saliency regions. Finally, it can improve the cognitive ability and decision-making ability for space image sequences. Therefore, we adopt the visual saliency technique to address the above key issue. This module contains the following three parts:

(1) Attention region computing: It focuses on the problem of the change of the attention scale and then captures the visual attention regions for space image sequences. The calculation process mainly includes two aspects: spatial attention regions and temporal attention regions.

(2) Frame rate conversion: It focuses on the problem of the change of the frame rate scale and improves the motion smoothness of space sequences. The calculation process mainly includes three steps: transform of attention blocks, acquisition of neighboring blocks and prediction of unknown pixels.

(3) Image resolution resizing: It focuses on the problem of the change of the resolution scale and improves the definition of space sequences with different resolutions under the guarantee of the global visual effects. The calculation process mainly includes two steps: partition interpolation computing and seam carving operation.

On the basis of transcale description, the module of perception-driven display is presented to achieve the task of high-quality spatial motion display under the condition of changing the different scales. And this module is an application of the first module and it could serve for the space navigation information processing system. This module also contains the following three parts:

(1) Target trajectory computing: It could adaptively display the motion trajectory of space targets according to the setting of scale parameters. The calculation process mainly includes five steps: attention region computing, frame rate conversion, boundary computing, trajectory coordinate calculation, and motion trajectory display.

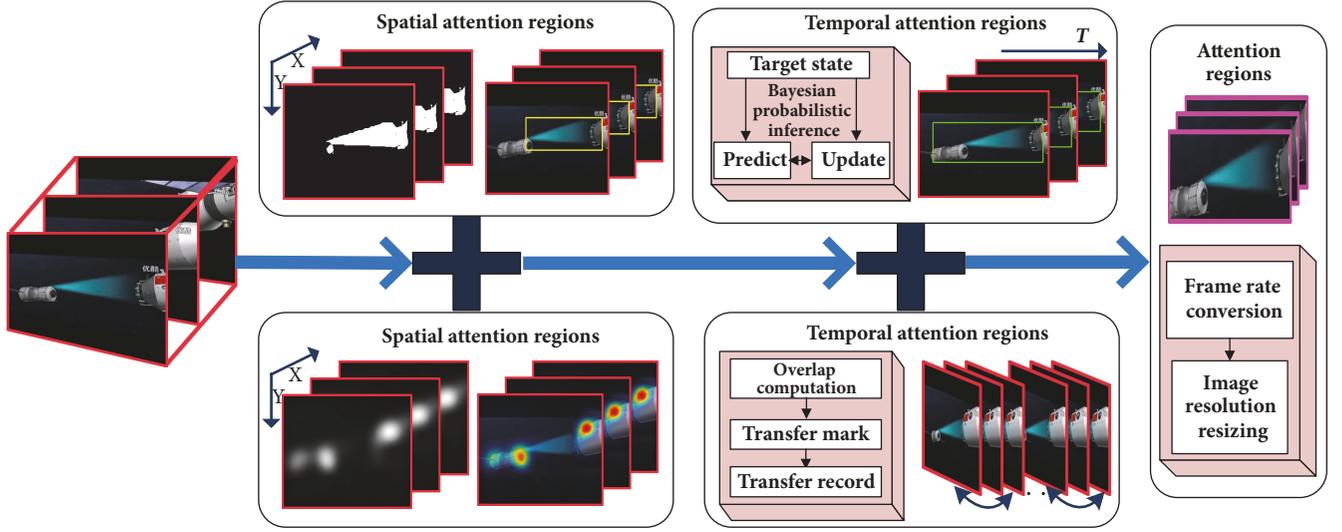


FIGURE 2: The flowchart of the module of attention region computing.

(2) Space transcale display. It obtains clear motion details under different time scales and spatial scales, and the calculation process mainly includes five steps: key frame computing, display map calculation, thumbnail computing, transcale display, and thumbnail resizing.

(3) Space movement display. It realizes the motion overview of the space targets in a narrative way. The calculation process mainly includes four steps: transition map setting, transition pixel calculation, transition region acquisition, and display map calculation.

### 3. The Module of Transcale Description

In this section, we will go over the individual parts of transcale description for space image sequences, including attention region computing, frame rate conversion, and image resolution resizing.

**3.1. Attention Region Computing.** The approach of attention region computing captures the important contents perceived by observers, i.e. visual attention regions, including spatial attention regions and temporal attention regions, and its flow chart is shown in Figure 2. Spatial attention regions are obtained according to the characteristics of color and image signature. And temporal attention regions are calculated by adopting visual tracking. Specifically, let  $I[1, n] = \{I_t\} (1 \leq t \leq n)$  be a space image sequence defined on the 3D space, and we use  $\{T, VA, KT\}$  to represent the attention region, where  $T$  denotes the collection of time stamp and  $VA$  denotes the attention regions containing spatial attention region  $VA_S$  and temporal attention region  $VA_T$ .  $KT$  denotes the transfer mark of attention regions. The computation of  $VA$  is shown in

$$VA(I_t) = \begin{cases} VA_S(I_t), & t = 1 \\ VA_S(I_t) \cup VA_T(I_t), & 1 < t \leq n \end{cases} \quad (1)$$

The computation of spatial attention regions is based on the research of biological vision; i.e., human's visual system is very sensitive to the contrast of the visual signal [12, 13]. We adopt the latest histogram color contrast method [14] to compute the spatial saliency value of any pixel in  $I_t$ , shown as follows:

$$Sals(\xi_{x_i, y_i}) = \frac{(\sum_{u=1}^{near} (G - D(\nu_c, \nu_u)) Sal(\xi_{x_i, y_i}))}{\sum_{u=1}^{near} (G - D(\nu_c, \nu_u))} \quad (2)$$

$$Sal(\xi_{x_i, y_i}) = \sum_{j=1, j \neq i}^{pn} D(\xi_{x_i, y_i}, \xi_{x_j, y_j}) = \sum_{\omega=1}^{cn} \phi_{\omega} D(\nu_c, \nu_{\omega}) \quad (3)$$

In (2) and (3),  $\xi_{x_i, y_i}$  and  $\xi_{x_j, y_j}$  denote any pixel in  $I_t$ ,  $Sals(\xi_{x_i, y_i})$  denotes the spatial saliency value of  $\xi_{x_i, y_i}$ ,  $D(\xi_{x_i, y_i}, \xi_{x_j, y_j})$  denotes the color distance function,  $p_n$  denotes the total number of pixels in  $I_t$ ,  $c_n$  denotes the number of distinct pixel colors in  $I_t$ ,  $\nu_c$  denotes the color value of  $\xi_{x_i, y_i}$ , and  $\phi_{\omega}$  denotes the probability of  $\nu_c$  in  $I_t$ . In addition,  $G = \sum_{u=1}^{near} D(\nu_c, \nu_u)$  in which  $near$  denotes the number of color that are nearest neighbors of  $\nu_c$ .

Then we quantify  $\xi_{x_j, y_j}$ , which has the larger value of spatial saliency, putting it into the set  $SAL^t = \{\xi_{x_j, y_j}\}$ . Because  $\xi_{x_j, y_j}$  may be in a very small, relatively unimportant region, so it is necessary to eliminate such a pixel. There are three elimination rules, shown as follows.

**Rule 1.** For  $\forall \varepsilon$ , if  $\xi_{x_j \pm \varepsilon, y_j} \notin SAL^t$  and  $\xi_{x_j, y_j \pm \varepsilon} \notin SAL^t$ , then  $\xi_{x_j, y_j}$  is eliminated from  $SAL^t$ , where  $\varepsilon = 1, 2, \dots$ .

**Rule 2.** For  $\forall \varepsilon'$  and  $\forall \varepsilon$ , if  $\xi_{x_j \pm \varepsilon', y_j} \in SAL^t$  and  $\xi_{x_j, y_j \pm \varepsilon} \notin SAL^t$ , then  $\xi_{x_j, y_j}$  is eliminated from  $SAL^t$ .

**Rule 3.** For  $\forall \varepsilon''$  and  $\forall \varepsilon$ , if  $\xi_{x_j, y_j \pm \varepsilon''} \in SAL^t$  and  $\xi_{x_j \pm \varepsilon, y_j} \notin SAL^t$ , then  $\xi_{x_j, y_j}$  is eliminated from  $SAL^t$ .

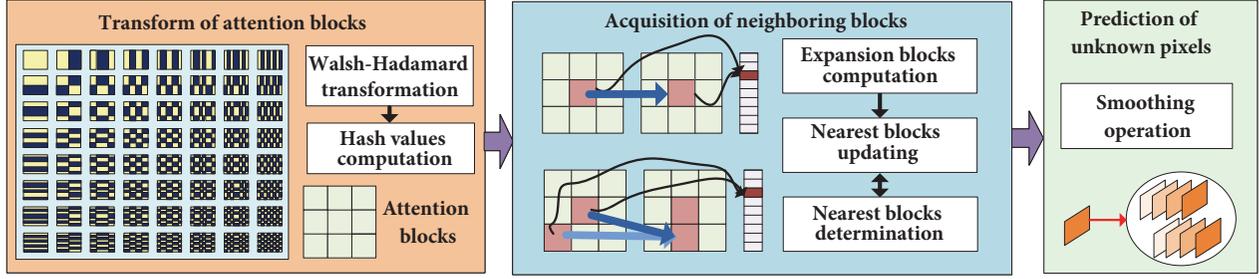


FIGURE 3: The computation of pixels in attention regions.

As  $\varepsilon' = 1, 2, \dots, (w_{initial} - j)$ ,  $\varepsilon'' = 1, 2, \dots, (h_{initial} - j)$ , and  $w_{initial}$  and  $h_{initial}$  are the initial resolution of image sequences. On this basis, the boundary values of  $VA_S(I_t)$  can be acquired by using the technology of bounding box.

The computation of temporal saliency regions is based on the observations of human attention continuity [15, 16]. According to the obtained spatial saliency regions, the temporal saliency regions are computed using visual tracking. Since the attention regions could vary over time, we use  $KT$  to describe the change extent of attention regions, shown as follows:

$$KT = \begin{cases} 1, & \frac{VA_S(I_t) \cap VA_S(I_{t+1})}{VA_S(I_t) \cup VA_S(I_{t+1})} > \delta_{SR} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

where  $\delta_{SR}$  denotes a threshold value. Then time  $t$ , which satisfies the above equation, is successively put into the set  $KTS$ . In our experiments, we determine the value  $\delta_{SR}$  to be 0.75.

We utilize Bayesian target tracking to compute  $VA_T(I_t)$ . Let  $\gamma_t = [x_t, y_t, dx_t, dy_t]^T$  be the state of  $VA_T(I_t)$ , in which  $[x_t, y_t]^T$  denotes the position vector and  $[dx_t, dy_t]^T$  denotes velocity vector. The set of spatial saliency region  $OS_{1:k} = \{VA_T(I_t)\}_{t=j_1 \dots j_2}$  is denoted as the observation set of  $VA_T(I_t)$ , in which  $j_1$  and  $j_2$  are the adjacent values in  $KTS$  and  $j_1 < j_2$ . Within the Bayesian inference framework, we estimate posterior probability function  $p(g_k | OS_{1:k})$  to compute target state  $g_k$  for  $VA_T(I_t)$ . The calculation of  $p(g_k | g_{k-1})$  includes prediction and updating, shown as follows:

$$p(g_k | OS_{1:k-1}) = \int p(g_k | g_{k-1}) p(g_{k-1} | OS_{1:k-1}) dg_{k-1} \quad (5)$$

$$p(g_k | OS_{1:k}) = \frac{p(OS_k | g_k) p(g_k | OS_{1:k-1})}{p(OS_k | OS_{k-1})} \quad (6)$$

where  $g_k$  and  $OS_k$  denote the state and observation value of  $k$ -th frame, respectively,  $p(g_k | g_{k-1})$  denotes the state transition model which can be solved using affine transformation, and  $p(OS_k | g_k)$  denotes state measurement model which can be computed using structural local sparse appearance method, shown as follows:

$$p(OS_k | g_k) = \sum_{j=1}^N (Dig(O))_k, \quad (7)$$

$$\chi_i = \frac{1}{M} \sum_{d=1}^k \sigma_i^{(d)}$$

where  $d$  denotes the patch divided by  $32 \times 32$  pixels in  $VA_S(I_m)$  (in which  $1 \leq m \leq k$ ),  $N$  denotes the number of the divided patches,  $\sigma_i$  denotes the sparse coding of  $d$ ,  $M$  denotes the normalization term, and  $\chi_i$  denotes the weighted sum of  $\sigma_i$ , setting set  $O$ . And  $Dig(\bullet)$  is the function of diagonal elements computing.

**3.2. Frame Rate Conversion.** To enhance motion smoothness of space image sequences, frame rate conversion is performed, whose core is to compute the high quality intermediate frames, i.e., how to calculate the pixel values of the intermediate frames. Here the partition method is used for calculation. Moreover, for non-attention regions, the values of pixels are unchanged. And for the attention regions, the values of pixels are accurately computed and the main computation steps contain transform of attention blocks, acquisition of neighboring blocks, and prediction of unknown pixels, which is shown in Figure 3. Specifically, attention regions in any two consecutive frames, i.e.,  $VA(I_t)$  and  $VA(I_{t+1})$ , are divided into  $\mu \times \mu$  overlapping image blocks, and each block is defined as an attention block,  $VAB$  for short. Note that we choose three different values for  $\mu$ , namely, 8, 16, and 32, and we find that the appropriate value is 16 according to the obtained experiment results. The set of  $VAB$  denotes as  $VAB^t = \{b_k^t\}$  and the computation procedure for the contained pixels is elaborated as follows.

Firstly, each  $VAB$  is projected using Walsh-Hadamard kernels [17], as shown below, and the results are stored in the temporary set  $WP = \{WP_j | 1 \leq j \leq \phi\}$ :

$$Q_{u1}^{(0)} = Q_{u2}^{(0)} = 1$$

$$Q_{u1}^{(k)} = \left\{ \left[ q_{u1}^{(k-1)} \rho_k q_{u1}^{(k-1)} \right] \right\}$$

$$Q_{u2}^{(k)} = \left\{ \left[ q_{u2}^{(k-1)} \rho_k q_{u2}^{(k-1)} \right] \right\}$$

$$Q_u^{(k)} = Q_{u1}^{(k)} \times Q_{u2}^{(k)}$$

$$\text{s.t. } q_{u1}^{(k-1)} \in Q_{u1}^{(k)},$$

```

Input: A spatial image sequence  $I$ 
Output: Changed image sequence  $I'$ 
(1) Compute the frame number of  $I$  as  $len$ .
(2) Give the first frame of  $I$  as the current frame  $I_1$ .
(3) Give the following frame of  $I_1$  as  $I_2$  and calculate their attention regions  $VA_1$  and  $VA_2$  according to the method in Section 3.1.
(4) Divide  $VA_1$  and  $VA_2$  into image block  $ST$  with  $\mu \times \mu$ .
(5) Compute the transformed value of each  $ST$  using equation (8) and create hash tables.
(6) FOR  $i = 1 \dots$  the number of hash tables
    FOR each transformed  $ST$  in  $VA_1$ 
        Obtain the expansion blocks according to spatial and temporal expansion rules;
        Compute the nearest blocks using the equation (9) and (10);
        Compute the value of pixels in  $VA_1$ ;
    END FOR
    END FOR
(7) Merge all the transformed  $ST$ .
(8) Compute interpolated frame  $I_1'$  by combining the attention region with non-attention region.
(9) Give the following frame of  $I_2$  as  $I_3, I_3 \rightarrow I_2$ .
(10) Repeat steps (3) to (9) until the frame number of the current frame is ( $len-1$ );
(11) Synthetic interpolated frames and original frames
Return  $I'$ 

```

ALGORITHM 1: Frame rate conversion algorithm.

$$\begin{aligned}
 q_{u2}^{(k-1)} &\in Q_{u2}^{(k)}, \\
 \rho_k &\in \{+1, -1\}
 \end{aligned} \tag{8}$$

Then the transformed result is assigned a hash value to accelerate VAB projection, setting in the corresponding hash table  $TB_m[h_m(b_z)] (1 \leq m \leq \phi)$ .

Secondly, we establish spatial and temporal expansion rules to obtain the expansion blocks for each VAB. For  $\forall b_k^t (b_k^t \in VAB^t)$ , each hash table  $TB_m$  needs to be operated according to expansion rules and the corresponding expansion block  $b_k^{t+1} (b_k^{t+1} \in VAB^{t+1})$  can be produced, putting into the matching set. On this basis, we determine the nearest blocks from the obtained expanding blocks using the freedom searching, shown as follows:

$$b_{near}^{n+1} = b_k^n + \omega \alpha^i R^i \tag{9}$$

In (9),  $R^i \in [-1, 1] \times [-1, 1]$ ,  $\alpha$  is a constant value and its value is 0.5,  $\omega$  is a search radius, and  $b_{near}^{n+1}$  is the best block of  $b_k^n$ . Then for each hash table, the following is used to compare and update the nearest blocks:

$$b_{near}^{n+1} = b_{map}^{n+1}, \quad \text{if } \text{disc}(b_k^n, b_{map}^{n+1}) < \text{disc}(b_k^n, b_{near}^{n+1}) \tag{10}$$

Finally, the predicted value of the pixels in the attention regions can be computed through a smoothing operation of the obtained nearest blocks (for detailed algorithm, see Algorithm 1).

**3.3. Image Resolution Resizing.** Image resizing has gained significant importance because of rapid increases in the diversity and versatility of display devices. However, existing methods do not resize results from the viewpoint of the

observers. In this section, we present a new resizing method to improve the resizing quality of the important contents perceived by observers. To accomplish this, two main issues need to be addressed. The first is the need to determine the important contents from the observers' viewpoint. The second is the need to improve the definitions of the important contents during resizing. For the first issue, we adopt the method of attention region computing (see Section 3.1) and, for the second one, we introduce a method of partition interpolation with an architecture that is illustrated in Figure 4, containing two steps: partition interpolation computing and seam carving operation. Specifically, we resize  $I_t$  from  $h_{initial} \times w_{initial}$  to  $h_{new} \times w_{new}$  and a resizing factor  $2^s$  is

$$\begin{aligned}
 s &= \min [s_H, s_W], \\
 s_H &= \left\lfloor \frac{h_{new}}{h_{initial}} \right\rfloor, \\
 s_W &= \left\lfloor \frac{w_{new}}{w_{initial}} \right\rfloor
 \end{aligned} \tag{11}$$

The idea of partition interpolation computing is to adopt different calculation approaches for different regions of original image sequences. Naturally, for any pixel  $i(x, y)$  in  $I_t$ , the interpolated pixel  $v_p$  is computed using

$$v_p = \begin{cases} AR_t'(2x+1, 2y+1), & i \in VA_t \\ BR_t'(2x+1, 2y+1), & i \in BR_t \end{cases} \tag{12}$$

where  $VA_t$  is the simplified symbol of  $VA(I_t)$ , which is computed in Section 3.1,  $BR_t$  denotes the general region (namely, nonattention region),  $VA_t \cup BR_t = I_t$ , and  $AR_t'(2x+1, 2y+1)$  and  $BR_t'(2x+1, 2y+1)$  denote the interpolated pixels in  $VA_t$  and  $BR_t$ , respectively.

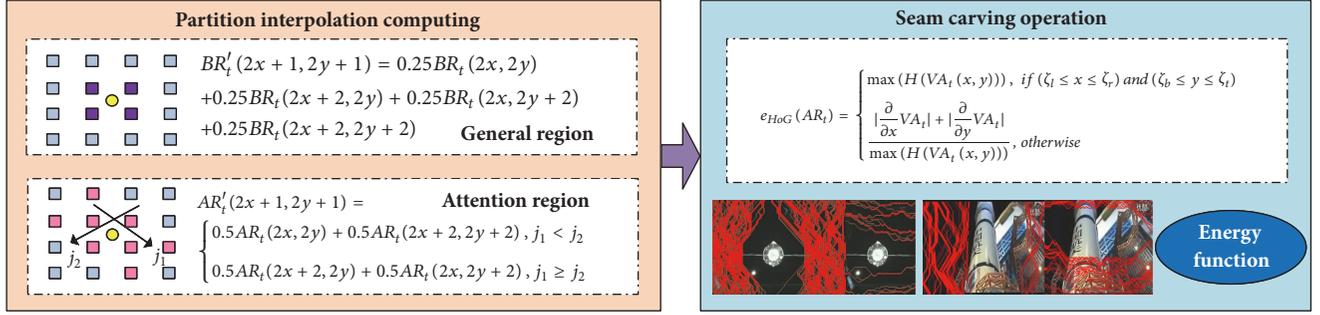


FIGURE 4: The method of partition interpolation.

The computation of  $AR_t''(2x+1,2y+1)$  contains two steps: initial estimation and energy modification. The initial value is calculated using

$$AR_t'(2x+1,2y+1) = \begin{cases} 0.5AR_t(2x,2y) + 0.5AR_t(2x+2,2y+2), & j_1 < j_2 \\ 0.5AR_t(2x+2,2y) + 0.5AR_t(2x,2y+2), & j_1 \geq j_2 \end{cases} \quad (13)$$

In (13),  $j_1$  and  $j_2$  are the determination function,  $j_1=Z_1-3Z_2+Z_3$ ,  $j_2=-3Z_1+Z_2+Z_4$ , and  $Z_i(i=1,2,3,4)$  is computed as follows:

$$Z_1 = AR_t(2x,2y) + AR_t(2x+2,2y+2),$$

$$Z_2 = AR_t(2x,2y+2) + AR_t(2x+2,2y)$$

$$Z_3 = AR_t(2x+4,2y) + AR_t(2x+2,2y-2)$$

$$Z_4 = AR_t(2x+4,2y+2) + AR_t(2x+2,2y+4) + AR_t(2x,2y-2) + AR_t(2x-2,2y) + AR_t(2x-2,2y+2) \quad (14)$$

For any pixel in general regions, the following is used to compute  $BR_t'(2x+1,2y+1)$ :

$$BR_t'(2x+1,2y+1) = 0.25BR_t(2x,2y) + 0.25BR_t(2x+2,2y) + 0.25BR_t(2x,2y+2) + 0.25BR_t(2x+2,2y+2) \quad (15)$$

Then we use seam carving method to preserve the contents of attention regions. That is to say, the pixels in the different regions are computed through energy function, shown as follows:

$$e_{HoG}(AR_t) = \begin{cases} \max(H(VA_t(x,y))), & \text{if } (\zeta_l \leq x \leq \zeta_r) \text{ and } (\zeta_b \leq y \leq \zeta_t) \\ \frac{|\partial/\partial x VA_t| + |\partial/\partial y VA_t|}{\max(H(VA_t(x,y)))}, & \text{otherwise} \end{cases} \quad (16)$$

where  $H(VA_t(x,y))$  represents the oriented gradient histogram of the pixels in the attention region and  $\zeta_l, \zeta_r, \zeta_b$ , and  $\zeta_t$  are the boundary values of  $VA_t$ .

#### 4. The Module of Perception-Driven Display

In this section, we describe in detail the formulation of our proposed the module of perception-driven display, containing three algorithms, namely, target trajectory computing, space transcale display, and space movement display. To some extent, all of them are an application of the module of transcale description and could be directly applied in the navigation information systems.

(1) *Target Trajectory Computing*. To display the motion trajectories of space targets under different scales, the approach of target trajectory computing is presented and it contains the following five steps:

*Step 1*. For each frame in  $I_t (1 \leq t \leq n)$ ,  $VA_t$  is computed by utilizing the method of attention region computing.

*Step 2*. Using the method of frame rate conversion,  $I[1, n]$  is changed into  $HI[1, N'] = HI_t(1 \leq t \leq N' \text{ and } N' > n)$  in which  $N$  is the number of changed frames and  $VA_t$  is updated in  $HI_t$ .

*Step 3.* The four boundary values are computed for each  $VA_t$ , including left boundary  $BV_L$ , right boundary  $BV_R$ , top boundary  $BV_T$ , and bottom boundary  $BV_B$ :

$$\begin{aligned} BV_L &= \min_{x_i \in VA_t} x_i, \\ BV_R &= \max_{x_i \in VA_t} x_i, \\ BV_T &= \min_{y_i \in VA_t} y_i, \\ BV_B &= \max_{y_i \in VA_t} y_i \end{aligned} \quad (17)$$

*Step 4.* For each  $VA_t$ , trajectory coordinates are calculated, including the value of horizontal coordinate  $VA_{row}$  and the value of vertical coordinate  $VA_{column}$ :

$$\begin{aligned} VA_{row} &= BV_L + \frac{(BV_R - BV_L)}{2}, \\ VA_{column} &= BV_T + \frac{(BV_B - BV_T)}{2} \end{aligned} \quad (18)$$

*Step 5.* Motion trajectory  $AT_v$  with the different scale  $v$  is computed using the following formula:

$$AT_v = \begin{cases} HI_1 \cup VA_1 \cup VA_2 \cup VA_3 \cup \dots \cup VA_N, & v = 1 \\ HI_1 \cup VA_1 \cup VA_3 \cup VA_5 \cup \dots \cup VA_{1+2 \cdot \lfloor (N-1)/2 \rfloor}, & v = 2 \\ \dots\dots\dots \\ HI_1 \cup VA_1 \cup VA_{1+\sigma} \cup VA_{1+2\sigma} \cup \dots \cup VA_{1+\sigma \cdot \lfloor (N-1)/2 \rfloor}, & v = \sigma \end{cases} \quad (19)$$

(2) *Space Transcale Display.* The aim of transcale display is not only to show motion pictures at different time scales but also to display with the different resolution scales. This algorithm mainly includes the following five steps:

*Step 1.* For each  $I_t$ , the key frame  $K_j$  ( $j = 1 \dots m$ ) is computed and then the corresponding attention region  $VAK_j$  is obtained.

*Step 2.* The transcale display map  $KM$  is captured by merging the attention regions under the condition of different key frames.

*Step 3.* For each  $VAK_j$  ( $j = 1 \dots m$ ), the corresponding thumbnail  $VU_j$  is computed by reducing the original image to the same resolution.

*Step 4.* The transcale display map  $KM'$  for  $I_t$  is computed:

$$KM' = KM \cup VU_1 \cup VU_2 \cup \dots \cup VU_m \quad (20)$$

*Step 5.* Using the method of image resolution resizing,  $VU_j$  could also have high-quality display with the different resolutions.

(3) *Space Movement Display.* To outline the motion process of space targets, a movement display algorithm is proposed, which focuses on smooth transition of space movement. Specifically, let  $K_p$  and  $K_q$  be the continuous key frames in  $I_t$  and the transition regions of them be  $KR_p(u, v)$  and  $KR_q(u', v')$ , respectively, in which  $u, u' \in [1, height]$ ,  $v \in [weight - r + 1, weight]$ ,  $v' \in [1, r]$ ,  $r$  denotes the transition radius, and *height* and *weight* are the current resolution of key frames. The process of this algorithm consists of the following four steps:

*Step 1.* Both  $KR_p$  and  $KR_q$  are resized to target resolution  $height \times 2r$ , and the transition region maps  $TMap_i$  ( $i = p, q$ ) are obtained.

*Step 2.* The value of pixel  $KR(u, v)$  in the transition region is computed:

$$KR(x, y) = c_1 \times TMap_p(x, y) + c_2 \times TMap_q(x, y) \quad (21)$$

in which  $c_1$  and  $c_2$  are transition parameters,  $c_1 = 1 - c_2$ ,  $c_2 = v/2r$ , and  $x$  and  $y$  are the position coordinates.

*Step 3.* Repeat Steps 1 and 2 until all the keyframe transition regions are calculated, and then new key frame  $K'$  is recorded.

*Step 4.* All the obtained key frames are merged to form display map  $MD$ :

$$MD = \sum_{j=1}^{\mu} K'_j \quad (22)$$

where  $\mu$  is the total number of key frames.

## 5. Experimental Results and Discussion

In this section, we conduct the quantitative and qualitative experiments to evaluate the performance of PTDS, including the transcale performance (see Sections 5.1, 5.2, and 5.3) and the display performance (see Sections 5.4, 5.5, and 5.6). We downloaded space video from Youku and they are segmented the video clips. Then we changed these clips to image sequences and formed our data shown in the Table 1. In all our experiments, we used the MATLAB platform and a PC with a 2.60 GHz Intel(R) Pentium(R) Dual-Core CPU processor with 1.96 GB of main memory.

*5.1. Evaluation of Attention Scale.* Figure 5 shows the obtained attention regions using our method, in which each

TABLE 1: Experimental data.

Sequence ID	Space image sequence	Initial resolution	Number of frames
S1	Aircraft automatic control	320×240	204
S2	Aircraft automatic homing	320×240	230
S3	Aircraft lift	640×480	238
S4	Aircraft docking	640×480	277
S5	Spacecraft detection	352×240	157
S6	Spacecraft launch	352×240	101
S7	Launch base	320×145	208
S8	Spacecraft flight	576×432	161
S9	Aircraft flight	1024×436	130

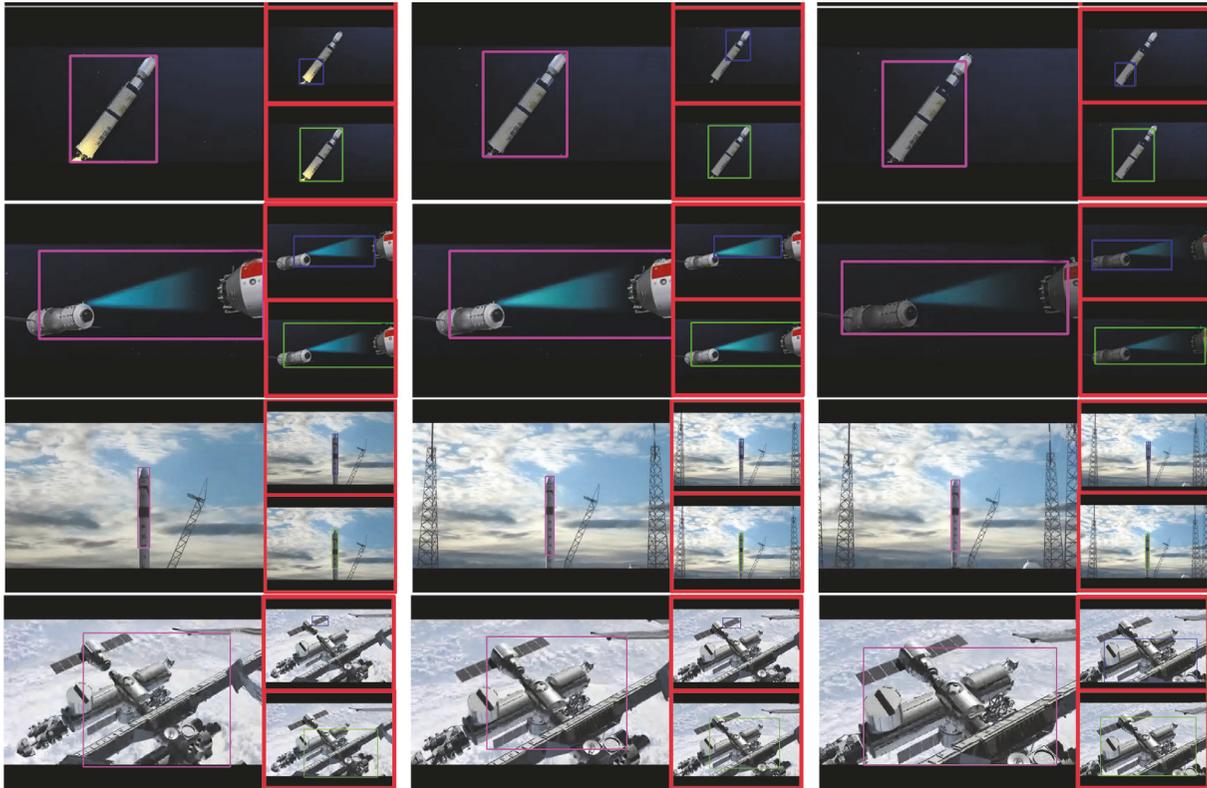


FIGURE 5: The obtained attention regions using PTDS from sample frames of different image sequences (from top to down, sequences S1, S2, S3, and S4). In each subfigure, the purple-bordered regions mark the attention regions, the blue-bordered regions mark the spatial attention regions, and the green-bordered regions mark the temporal attention regions.

row exhibits the results of different space image sequences and each column exhibits sample frames of each sequences. In addition, each subfigure consists of three parts: the purple-bordered region in the left marks the attention regions, the blue-bordered region in the upper right marks the spatial attention region, and the green-bordered region in the lower right marks the temporal attention region. Specifically, the first row shows the results for frames 100, 128, and 134 from S1, and the second row shows the results for frames 119, 121, and 135 from S2. Observing these rows, we can see that the spatial attention regions only capture part of space targets, for example, the result in the first row and first column. This means that both the calculations of spatial attention regions

and temporal attention regions are necessary for obtaining the accurate attention regions, especially for image sequences with large motion amplitude. Moreover, the third row shows the results for frame 10, 24 and 30 from S3, and the last row shows the results for frames 8, 45, and 88 from S4. These sequences have a relatively complicated background, which causes great interference to the detection of attention regions. However, the proposed method still obtains good results and these results lay the foundation for transcale display.

Figure 6 shows the quantitative comparison results for the other methods, including image signature detection (ISD) [18] and global contrast detection (GCD) [14]. Here we adopt

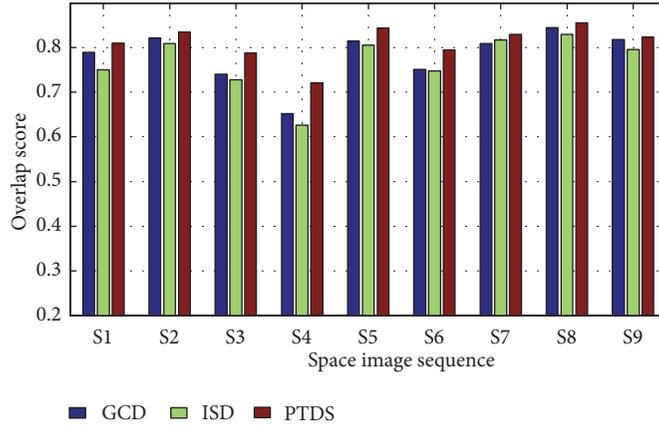


FIGURE 6: Comparison results of the average overlap score.

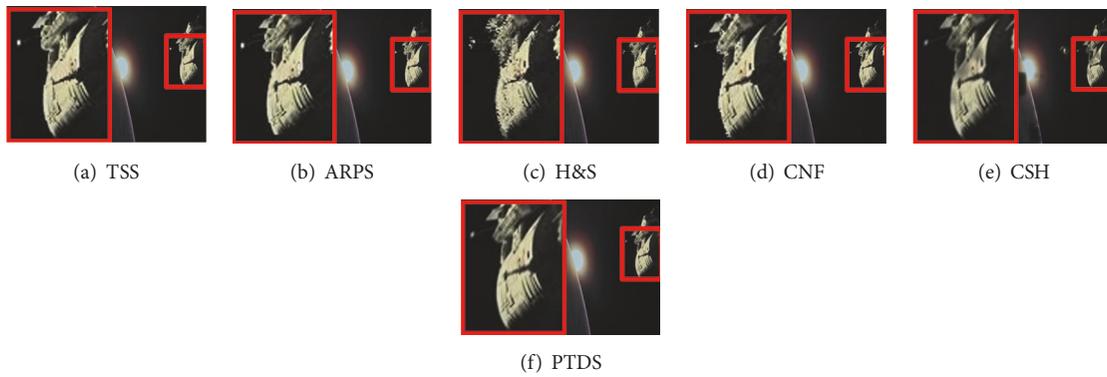


FIGURE 7: Interpolation frame 38 of sequence S5.

the overlap score [19] as a quantitative indicator, which is defined as the following formula:

$$score = \frac{R_S \cap R_G}{R_S \cup R_G} \quad (23)$$

where  $R_S$  denotes the obtained attention result and  $R_G$  denotes the corresponding ground truth bounding box. The larger the overlap scores are, the more accurate the results are. From this figure, we can see the average overlap scores using three methods. It is noted that our method achieves a larger score than those of the other methods, and the results show that the obtained attention regions using our method are closer to the true significant region of original image sequences.

**5.2. Evaluation of Frame Rate Scale.** The quality of interpolated frames is a key factor to evaluate the change of the scale of frame rate, and in experiments, we removed one in every two consecutive frames of the original image sequence and reconstructed the removed one adopting the different methods, including three-step search (TSS), adaptive rood pattern search (ARPS), Horn and Schunck (H&S), CNF [20], CSH [21] and our proposed method. Figures 7 and 8 show the frame interpolation results for sequence S5 and S6, and from the red-bordered region of each subfigure, the visual differences among algorithms can be determined. TSS and

ARPS exhibit a poor interpolation effect, H&S and CNF introduce a suspension effect, CSH produced disappointing effects, and PTDS shows the clear details of spacecraft. From these figures, it is evident that the proposed method shows comparatively better performance in terms of visual quality.

Figure 9 shows the average quantitative values of each image sequence by using the different methods. The left part demonstrates the average PSNR values, which is traditional quantitative measure in the term of accuracy, and the right part demonstrates the MSSIM results, which assess the image visibility quality from the viewpoint of image formation, under the assumption of a correlation between human visual perception and image structural information. In this figure, it is obvious that our method could achieve the highest average values.

**5.3. Evaluation of Resolution Scale.** We resize the original image sequences by using the different methods, including scaling, the best cropping, improved seam carving (ISC) [22], and our method, and then the obtained results are compared to evaluate the performance on the change of resolution scales.

Figure 10 shows the comparison results of sequence S7 from  $320 \times 145$  to  $500 \times 230$ . Figure 10(a) shows the original frames, in which the numbers of frames are 131, 138, 147, 151, 174, 179, 189, 203, and 208. Figure 10(b) shows that using

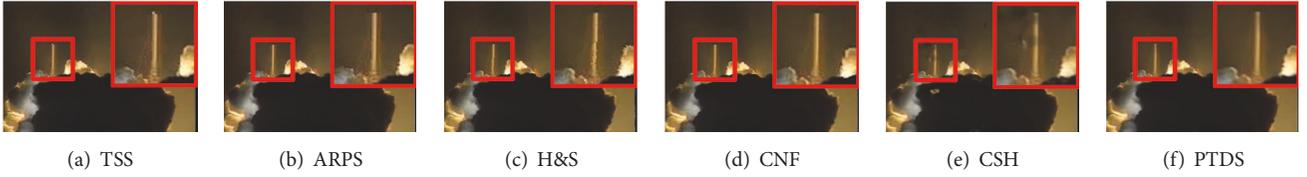


FIGURE 8: Interpolation frame 4 of sequence S6.

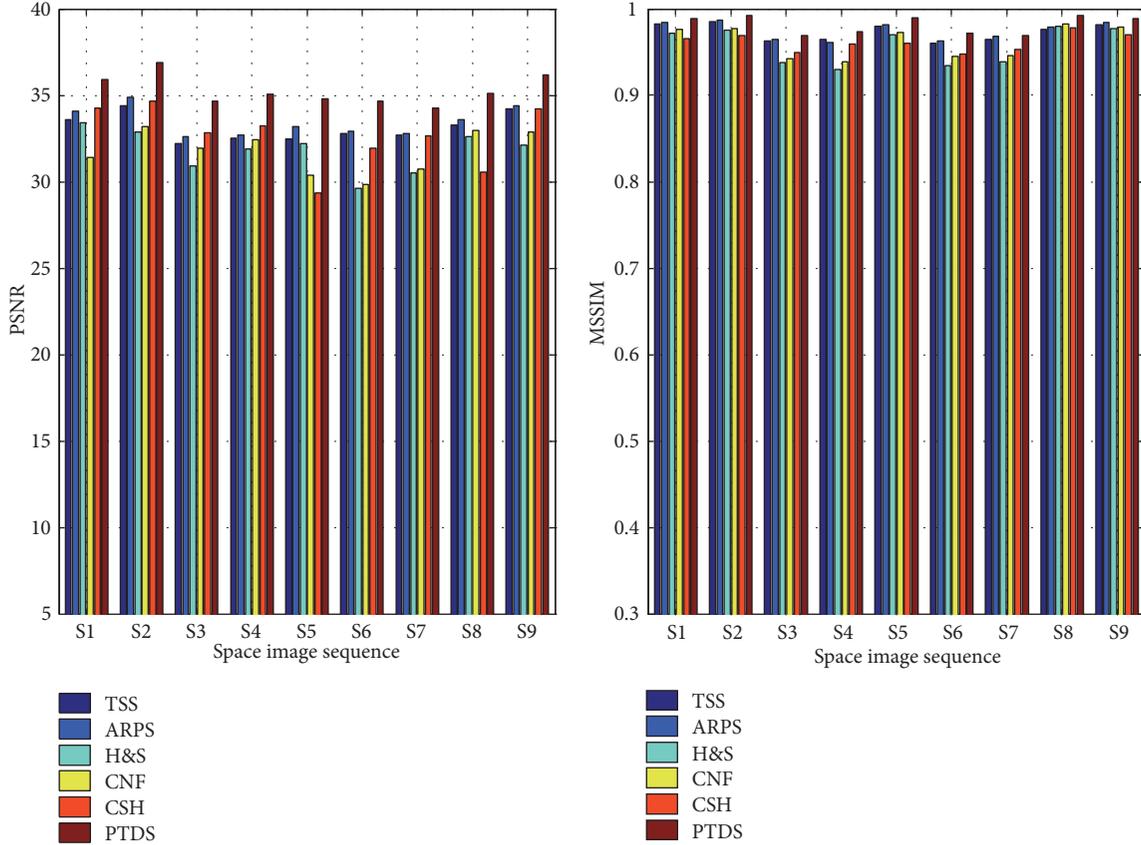


FIGURE 9: Comparison results of average values in the term of PSNR and MSSIM.

scaling method, the launch base becomes vaguer than before. Figure 10(c) shows that using the best cropping method, the launch base is only partly displayed, resulting in the original information becoming missing. Figure 10(d) shows that using the improved seam carving method, the prominent part of launch base becomes smaller than before, indicating that ISC is not suitable for image enlarging. Figure 10(e) shows that our method clearly displays the prominent objects of the original frames and ensures a global visual effect when the resolution scales are changed. Similarly, resizing results of the sequence S8 are shown in Figure 11.

**5.4. Evaluation of Target Trajectory Computing.** Figure 12 shows target trajectory of sequence S2 at scales  $v = 1$ ,  $v = 10$ , and  $v = 20$ . As can be seen from this figure, as the value of  $v$  increases, the movement amplitude of space targets becomes greater and the description of the trajectory could be coarser, which clearly characterizes the change of the movement of

automatic homing. Similarly, target trajectory of sequence S5 is shown in Figure 13.

Figure 14 shows the detailed positions of target trajectories corresponding to Figure 12, which are described in way of computing the position of the center pixel in attention regions. And it can be seen that the larger the scale  $v$  is, the fewer the description points are and the rougher the position trajectory of center point is and vice versa. Similarly, Figure 15 shows target trajectories corresponding to Figure 13. In short, PTDS exhibits good performance for displaying space target trajectories under different scales, which can lay a foundation for space tasks, such as condition monitoring and motion tracking.

**5.5. Evaluation of Space Transcale Display.** Figure 16 gives transcale displays for sequences S5 and S9, in which the attention regions are shown in the left part of each subfigure. Figure 16(a) shows both the whole process of detecting from

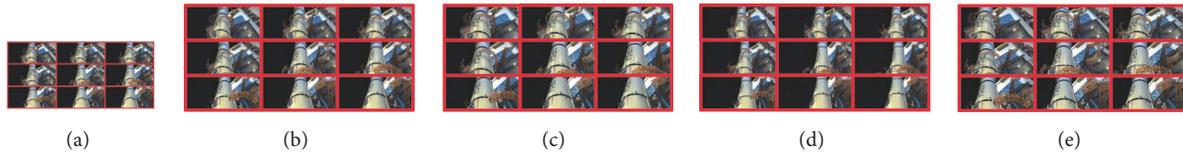


FIGURE 10: Comparison results of sequence S7 using four methods when the resolution is resized from  $320 \times 145$  to  $500 \times 230$ . (a) The original frames, (b) scaling, (c) the best cropping, (d) ISC, and (e) PTDS.

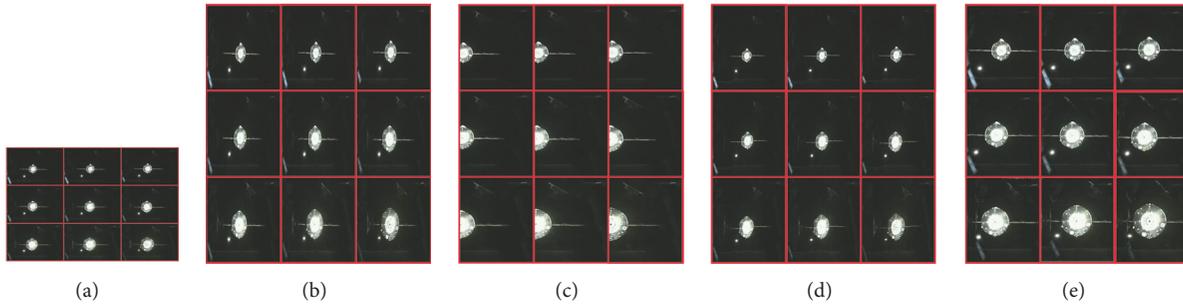


FIGURE 11: Comparison results of sequence S8 using four methods when the resolution is resized from  $576 \times 432$  to  $600 \times 800$ . (a) The original frames, (b) scaling, (c) the best cropping, (d) ISC, and (e) PTDS.

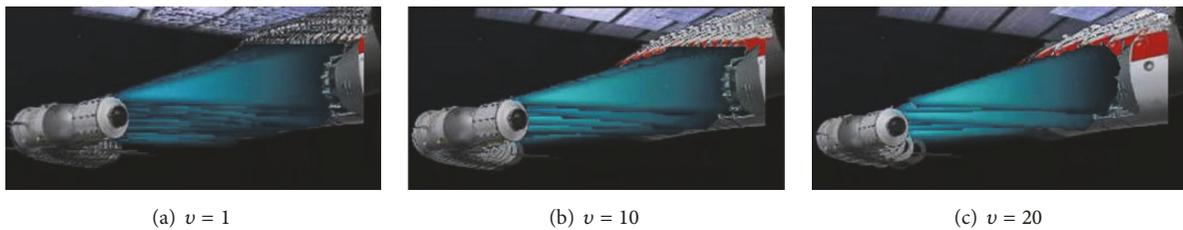


FIGURE 12: Target trajectory display of sequence S2.

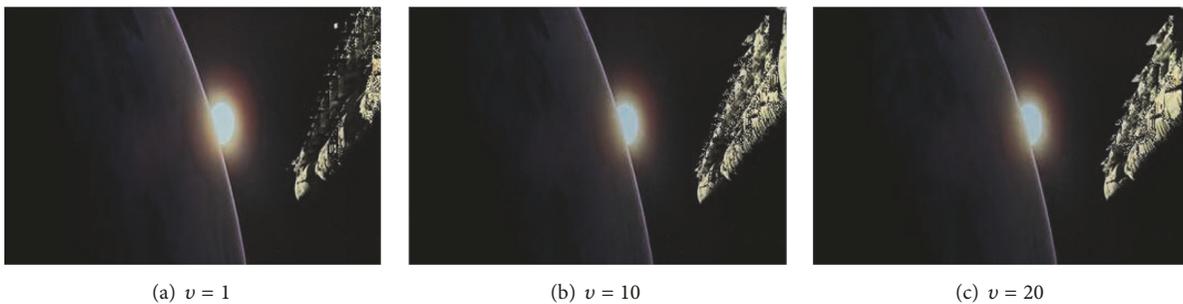


FIGURE 13: Target trajectory display of sequence S5.

near to far and the motion details of the space targets. Similarly, the process of movement from far to near is given in Figure 16(b). From this figure, we can see that the proposed space transcale display algorithm can clearly exhibit the motion process of the space targets at different scales. And by observing the accompanying attention regions, motion details of space targets, such as target pose and operating status, are displayed at different distances and time, which fully reflects the transcale characteristics of the PTDS scheme.

Figure 17 summarizes the objective evaluation values of space transcale display corresponding to Figure 16, including average gradient, edge intensity, spatial frequency, image

definition, and quality score. With the increase of the number of key frames, the values of the above evaluation index are also increasing, indicating that the space transcale display algorithm can obtain high-quality transcale display map.

5.6. *Evaluation of Space Movement Display.* Figure 18 shows Space movement display for space image sequences, in which a, b, and c are, respectively, S3, S4, and S6. By adopting the narrative diagram, the complete movement of the space targets are clearly exhibited. Figures 18(b), 18(d), and 18(f) are their corresponding transition area details, depicting the comparison results using direct fusion method in the left and

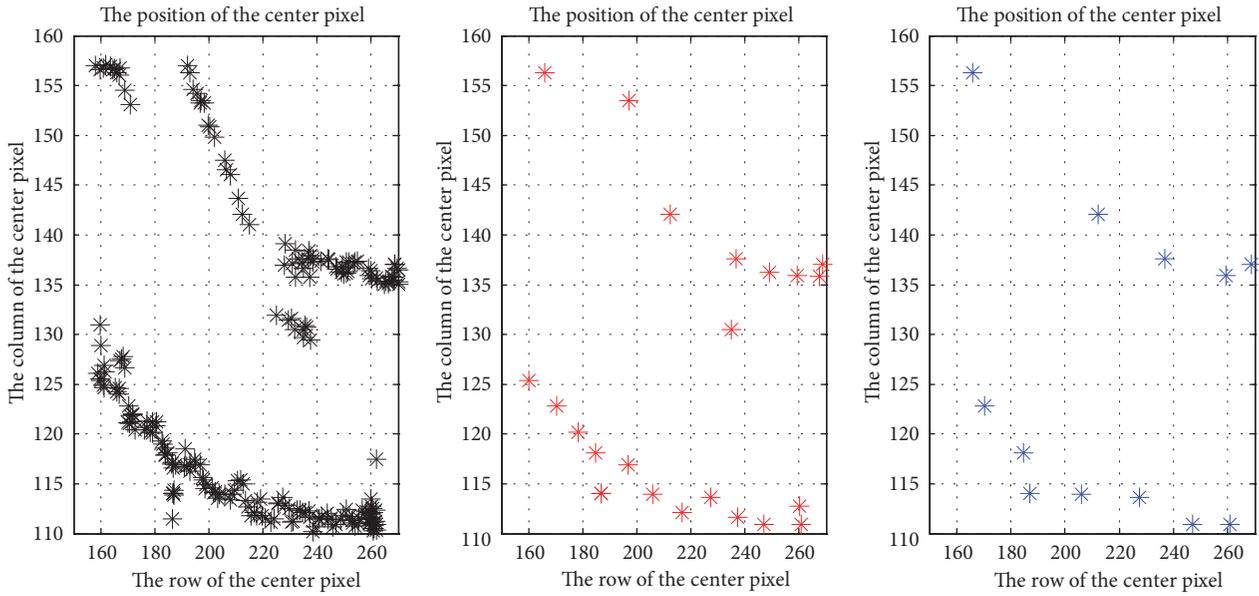


FIGURE 14: Target trajectory of the center pixel corresponding to sequence S2. From left to right,  $v = 1$ ,  $v = 10$ , and  $v = 20$ .

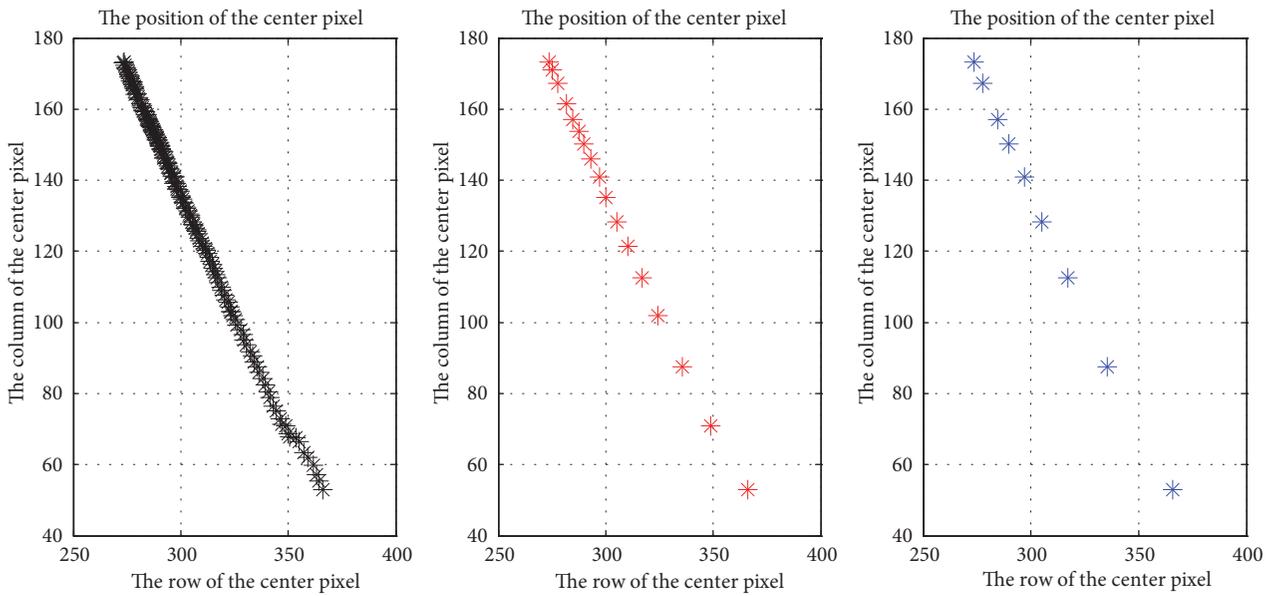
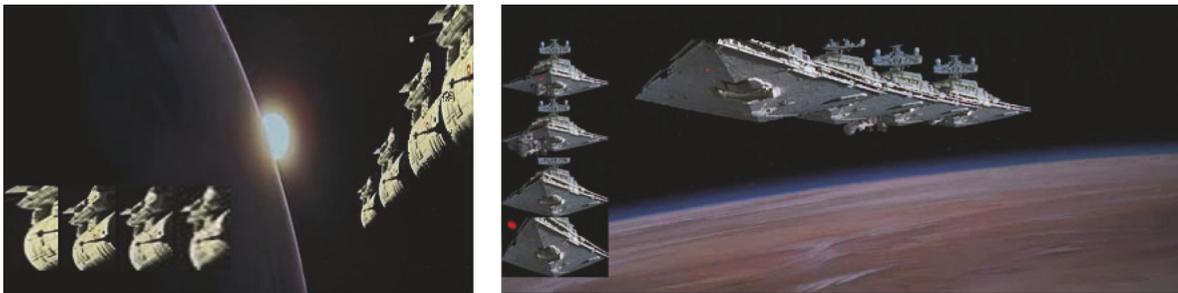


FIGURE 15: Target trajectory of the center pixel corresponding to sequence S5. From left to right,  $v = 1$ ,  $v = 10$ , and  $v = 20$ .



(a) S5

(b) S9

FIGURE 16: Space transcale display of space image sequences.

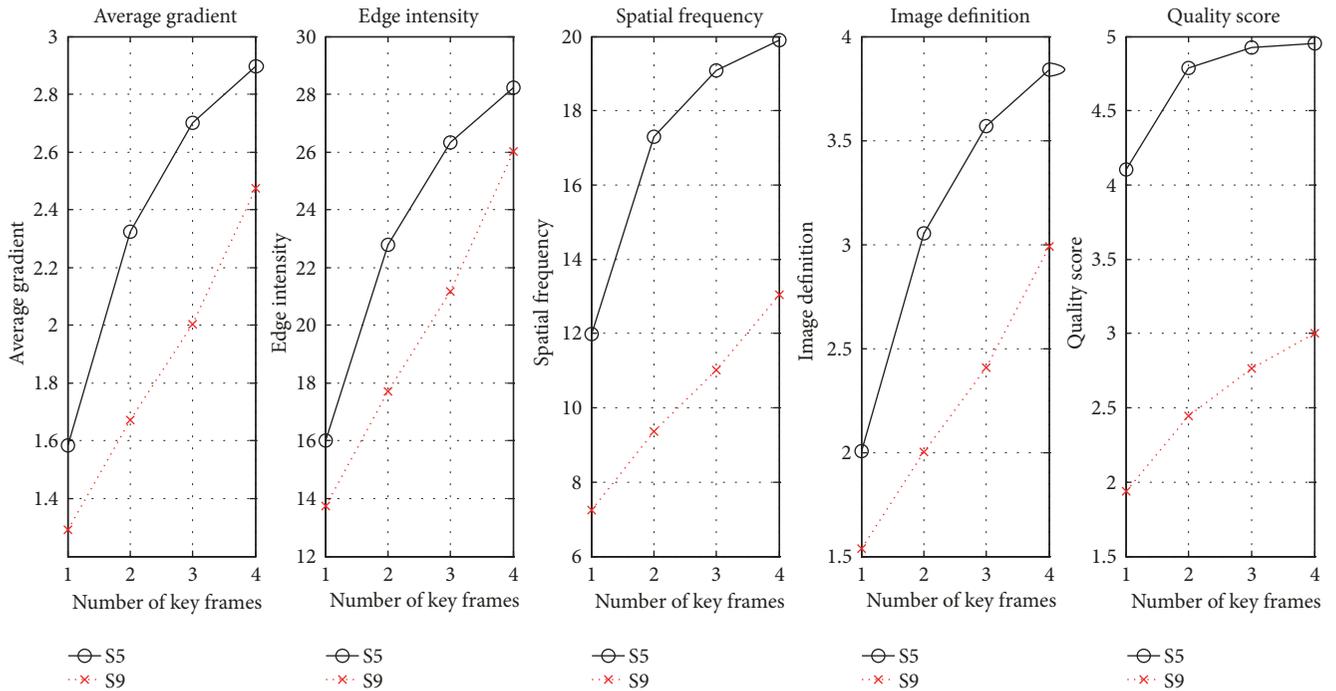


FIGURE 17: Objective evaluation.

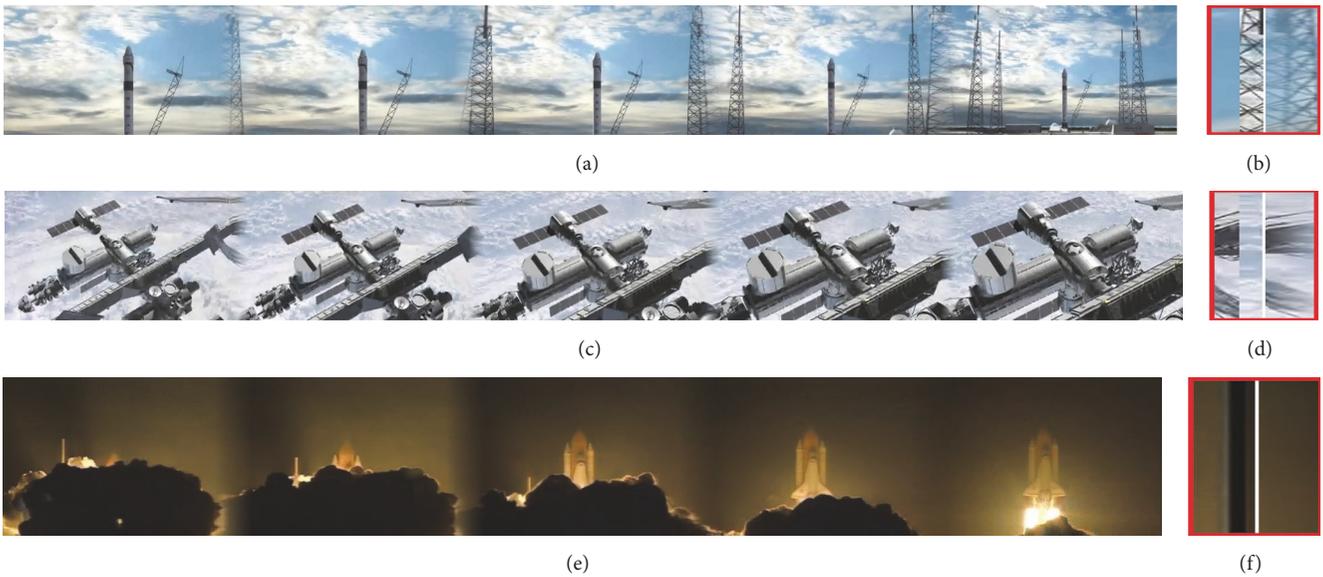


FIGURE 18: Space movement display for S3, S4, and S6.

our method in the right. For the left part, it is obvious to see a straight line to distinguish between two key frames, while for the right part, smooth and delicate transition areas are displayed to achieve a perfect representation of the movement process.

### 6. Conclusions

In this paper, we focus on the transcale problem of spatial image sequences and propose a novel display scheme which improves the awareness of multimedia processing.

The contribution of this scheme manifests in two aspects. On the one hand, space targets sustained attention by using visual saliency technology and then the details of the movement of the targets are enhanced. On the other hand, the motion of spatial image can be smoothly reproduced and effectively showed on the display device with different sizes and resolutions. Experimental results show that the proposed method outperforms the representative methods in terms of image visualization and quantitative measures. In the future, we will further study the transcale characteristics of space image sequences and explore the mutual influence between

the different scales. On this basis, we will also establish a robust transcale processing mechanism to better serve spatial rendezvous and docking.

## Data Availability

The data used to support the findings of this study have not been made available because the data is confidential.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (61702241, 61602227), the Foundation of the Education Department of Liaoning Province (LJYL019), and the Doctoral Starting Up Foundation of Science Project of Liaoning Province (201601365).

## References

- [1] W. Wang and Y. N. Hu, "Accuracy performance evaluation of Beidou navigation satellite system," *Acta Astronomica Sinica*, vol. 58, no. 2, 2017.
- [2] C. Shi, Q. Zhao, M. Li et al., "Precise orbit determination of Beidou Satellites with precise positioning," *Science China Earth Sciences*, vol. 55, no. 7, pp. 1079–1086, 2012.
- [3] J. Luo, S. Wu, S. Xu, J. Jiao, and Q. Zhang, "A cross-layer image transmission scheme for deep space exploration," in *Proceedings of the 86th Vehicular Technology Conference (VTC-Fall '17)*, pp. 1–5, IEEE, September 2017.
- [4] P. O'Driscoll, E. Merényi, and R. Grossman, "Using spatial characteristics to aid automation of SOM segmentation of functional image data," in *Proceedings of the 12th International Workshop on Self-Organizing Maps and Learning Vector Quantization, Clustering and Data Visualization (WSOM)*, pp. 1–8, June 2017.
- [5] D. Meng, Y. Jia, K. Cai, and J. Du, "Transcale average consensus of directed multi-vehicle networks with fixed and switching topologies," *International Journal of Control*, vol. 90, no. 10, pp. 2098–2110, 2017.
- [6] L. Zhao and Y. Jia, "Transcale control for a class of discrete stochastic systems based on wavelet packet decomposition," *Information Sciences*, vol. 296, no. 1, pp. 25–41, 2015.
- [7] L. Zhao, Y. Jia, J. Yu, and J. Du, " $H_\infty$  sliding mode based scaled consensus control for linear multi-agent systems with disturbances," *Applied Mathematics and Computation*, vol. 292, pp. 375–389, 2017.
- [8] B. J. White, D. J. Berg, J. Y. Kan, R. A. Marino, L. Itti, and D. P. Munoz, "Superior colliculus neurons encode a visual saliency map during free viewing of natural dynamic video," *Nature Communications*, vol. 8, article 14263, 2017.
- [9] J. Yang and M.-H. Yang, "Top-down visual saliency via joint CRF and dictionary learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 3, pp. 576–588, 2017.
- [10] S. Bhattacharya, K. S. Venkatesh, and S. Gupta, "Visual saliency detection using spatiotemporal decomposition," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 1665–1675, 2018.
- [11] V. Ramanishka, A. Das, J. Zhang, and K. Saenko, "Top-down visual saliency guided by captions," in *Proceedings of the 30th Conference on Computer Vision and Pattern Recognition, CVPR '17*, IEEE, 2017.
- [12] S. I. Cho, S.-J. Kang, and Y. H. Kim, "Human perception-based image segmentation using optimising of colour quantisation," *IET Image Processing*, vol. 8, no. 12, pp. 761–770, 2014.
- [13] M. Eickenberg, A. Gramfort, G. Varoquaux, and B. Thirion, "Seeing it all: convolutional network layers map the function of the human visual system," *NeuroImage*, vol. 152, pp. 184–194, 2017.
- [14] M. M. Cheng, G. X. Zhang, N. J. Mitra, X. Huang, and S. Hu, "Global contrast based salient region detection," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11)*, pp. 409–416, Providence, RI, USA, June 2011.
- [15] X. Jia, H. Lu, and M. Yang, "Visual tracking via adaptive structural local sparse appearance model," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '12)*, pp. 1822–1829, June 2012.
- [16] G. Mehraei, B. Shinn-Cunningham, and T. Dau, "Influence of spatial and non-spatial feature continuity on cortical alpha oscillations," *The Journal of the Acoustical Society of America*, vol. 141, no. 5, pp. 3634–3635, 2017.
- [17] M. T. Hamood and S. Boussakta, "Fast Walsh-Hadamard-Fourier transform algorithm," *IEEE Transactions on Signal Processing*, vol. 59, no. 11, pp. 5627–5631, 2011.
- [18] X. Hou, J. Harel, and C. Koch, "Image signature: highlighting sparse salient regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 194–201, 2012.
- [19] M. Everingham, L. van Gool, C. K. I. Williams, J. Winn, and A. Zisserman, "The pascal visual object classes (VOC) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2010.
- [20] D. Sun, S. Roth, and M. J. Black, "Secrets of optical flow estimation and their principles," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '10)*, pp. 2432–2439, IEEE, California, Calif, USA, June 2010.
- [21] S. Korman and S. Avidan, "Coherency sensitive hashing," in *Proceedings of the 2011 IEEE International Conference on Computer Vision, ICCV 2011*, pp. 1607–1614, Spain, November 2011.
- [22] D. D. Conger, M. Kumar, R. L. Miller, J. Luo, and H. Radha, "Improved seam carving for image resizing," in *Proceedings of the 2010 IEEE Workshop on Signal Processing Systems, SiPS 2010*, pp. 345–349, USA, October 2010.



**Hindawi**

Submit your manuscripts at  
[www.hindawi.com](http://www.hindawi.com)

