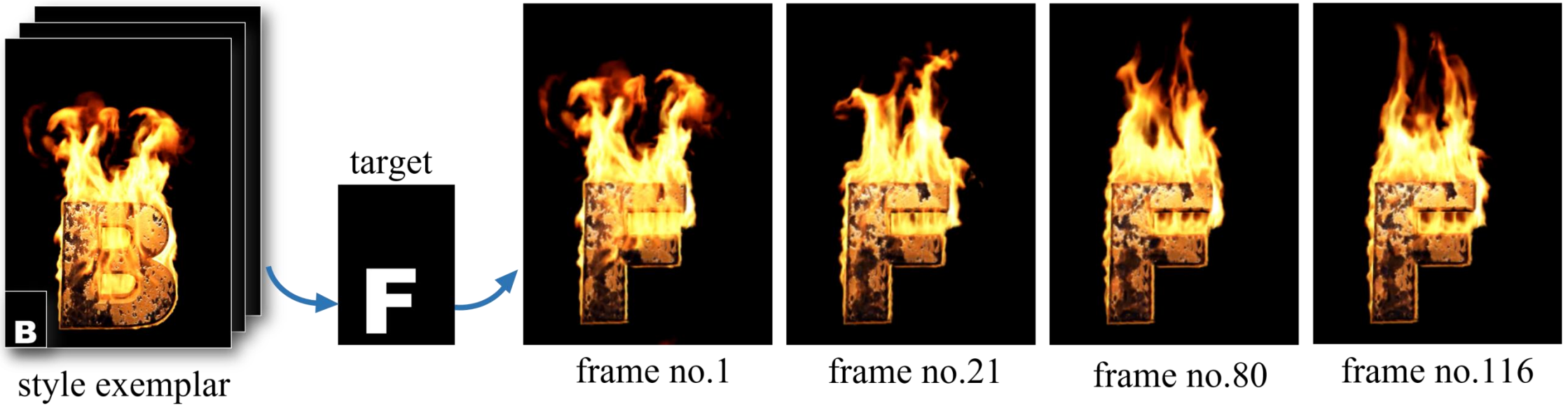
北京大学
PEKING UNIVERSITY

DynTypo: Example-based Dynamic Text Effects Transfer



连宙辉

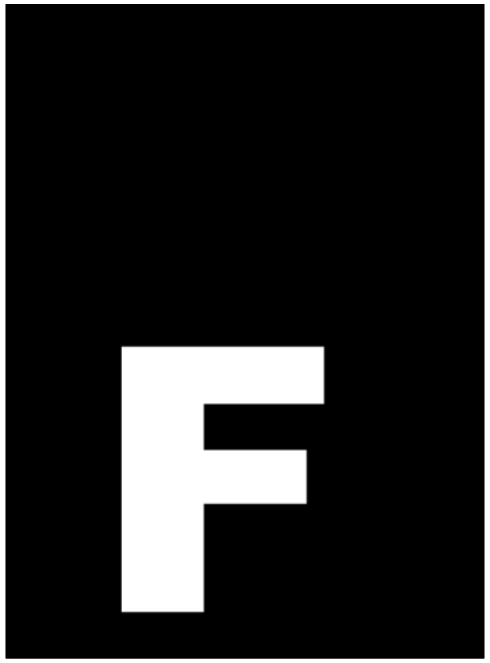
Yifang Men, *Zhouhui Lian**, Yingmin Tang, Jianguo Xiao

Institute of Computer Science and Technology, Peking University

Motivation

Animate a still image of the target text by transferring the desired dynamic effects from an observed exemplar

Exemplar



Target text

Motivation

Animate a still image of the target text by transferring the desired dynamic effects from an observed exemplar

Exemplar

Stylized result

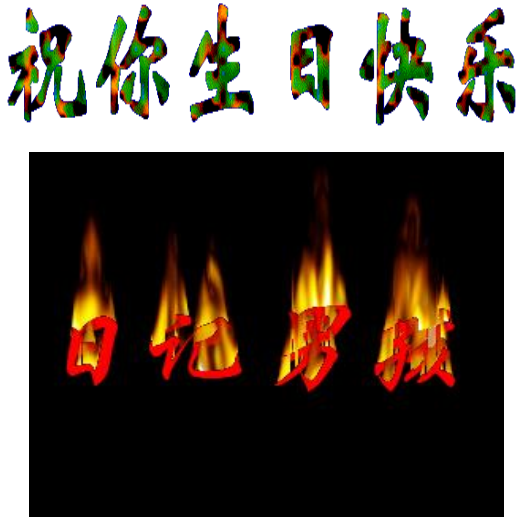
Motivation

Not only challenging in academic researches
but also of great commercial values in practice

Widely used in the movies, advertisement and video clips

Difficult to produce manually

- skill requirements
- time consuming



Common plain effects in daily life

Motivation

Not only challenging in academic researches
but also of great commercial values in practice

The native extension of applying text effects transfer [Yang et al. 2017] to consecutive frames independently

- Temporal Artifacts — produce strong flickers due to subtle appearance changes
- Appearance Artifacts — fail to transfer intense flame effects

Motivation

Not only challenging in academic researches
but also of great commercial values in practice

Text effects transfer [Yang et al. 2017] + Common temporal constraints [Fiser et al. 2017]

It improves temporal smoothing but still generates subtle trembling for static textures.

Motivation

Main Challenges

- We use only **a single image of the target text** as input (not target video or a sequence of motion fields with flow guidance) to animate and stylize it.
- **The composition of effect patterns** is more **complicated** with the static and dynamic effects often blended.

Temporal Artifacts

- **Little semantic information** is contained in raw text images and no guidance for intense text effects.

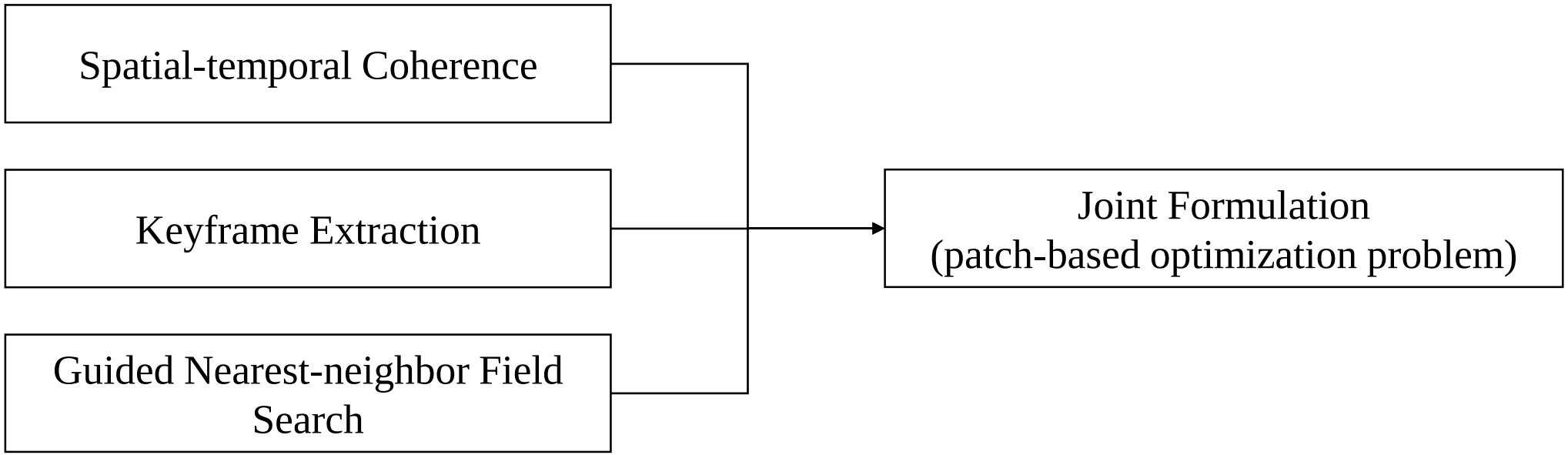
Appearance Artifacts

Method

Problem Formulation and Analysis

Our Approach

Overview



Method

1. Joint Formulation

Combine the following implementations into one joint patch-based optimization problem

Energy function

$$E = \sum_{Q \in T} \min_{P \in S} (\lambda D(P_{text}, Q_{text}) + \quad D(P^t_{sty}, Q^t_{sty}))$$

Optimized with our Guided NNF Search

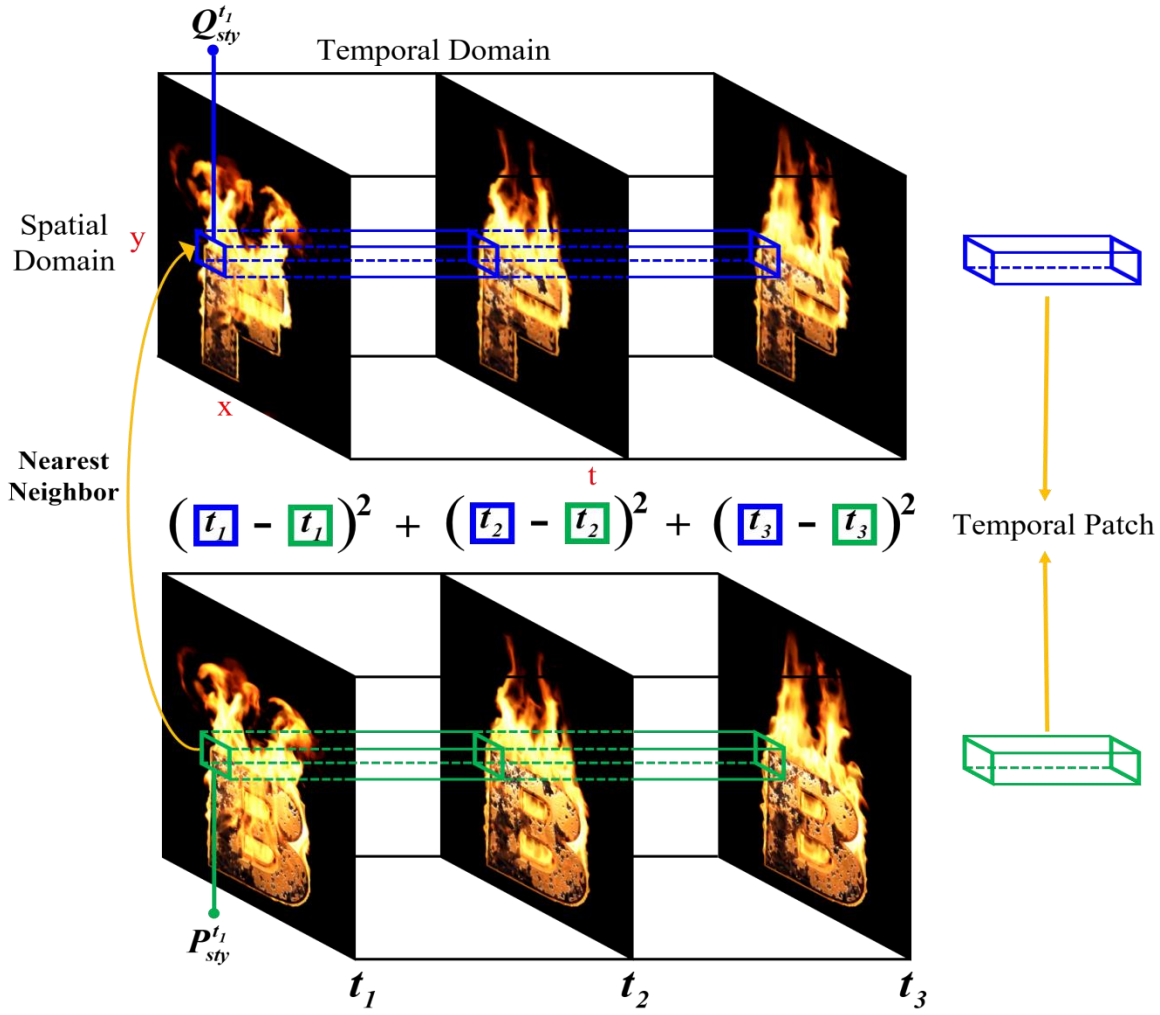
A common NNF

Applied to all frames

Target stylized video

Method

2. Spatial-temporal Coherence



- **A common NNF** for **temporal consistency**
- **Keyframes** for **spatial continuity**

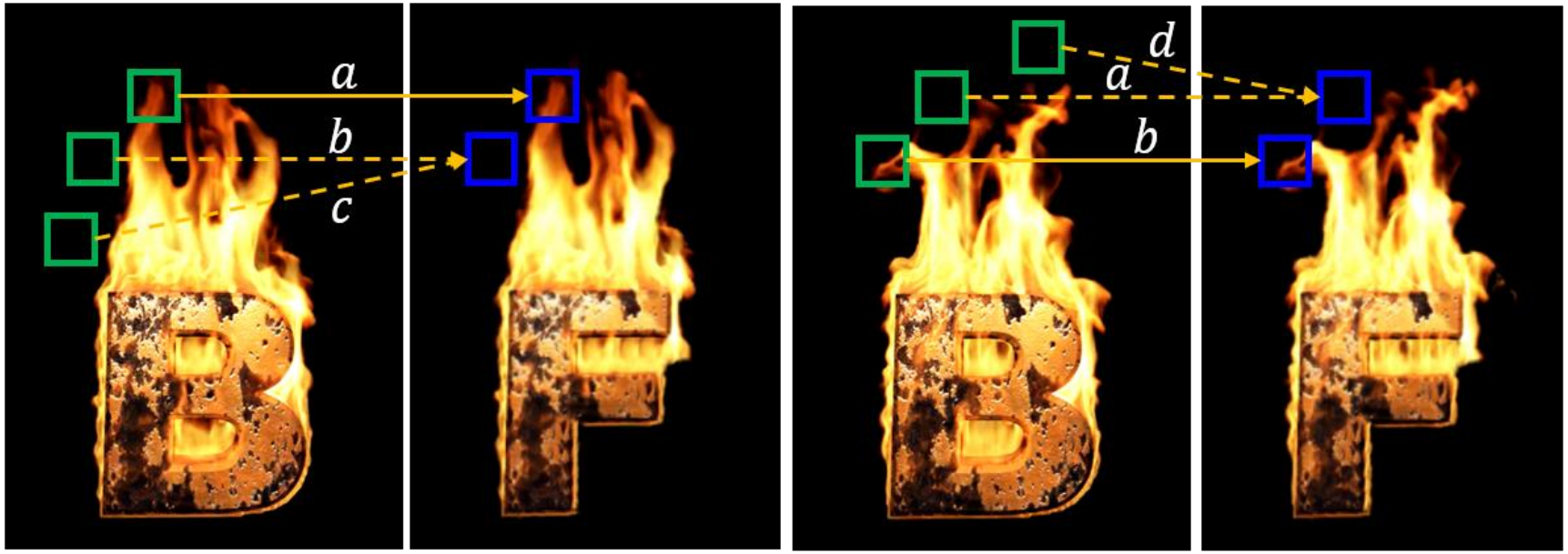
Spatial-temporal coherence term

$$E_{st} = \sum_{t \in kf} D(P^t_{sty} - Q^t_{sty})$$

Method

3. Keyframe Extraction

Introduce more constraints for texture coherence with keyframes containing more representative textural features.



Method

3. Keyframe Extraction

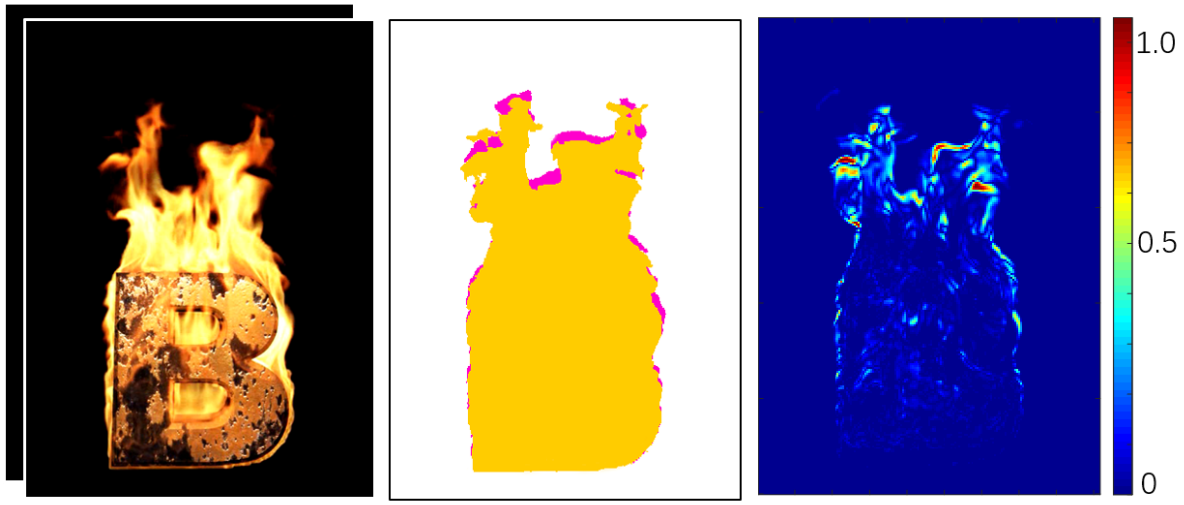
The keyframes are expected to **show violent movements** compared with **previous frames**, especially more emitters-places where new fluid is spawned (pixels marked in pink)

The map of color-changed values at t^{th} frame

$$m_t = | g(S_{sty}^{t}) - g(S_{sty}^{t-1}) |, \ 2 \le t \le N$$

The value of motion intensity at t^{th} frame

$$v_t = \sum_{p=1 \ldots w \times h} \tilde{m}_t(p)$$



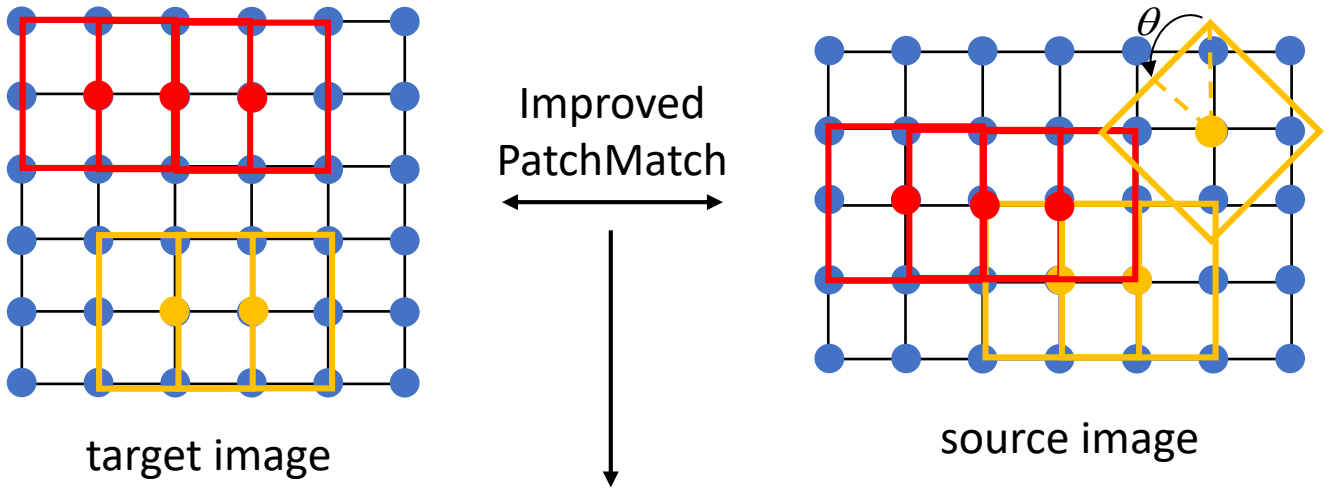
(b) $S_{sty}^{t-1,t}$ (c) $M_{sty}^{t-1,t}$ (d) $\tilde{m}_t$

Method

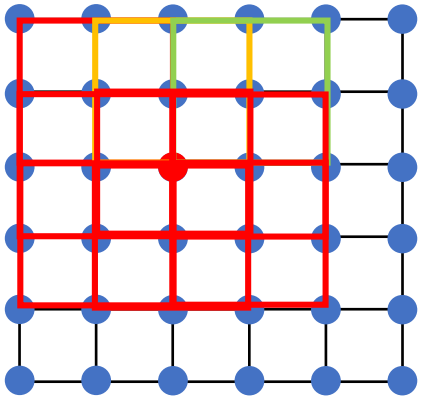
4. Guided Nearest-neighbor Field Search

PatchMatch Algorithm

1. Search — find nearest-neighbor field



2. Vote — reconstruct target image

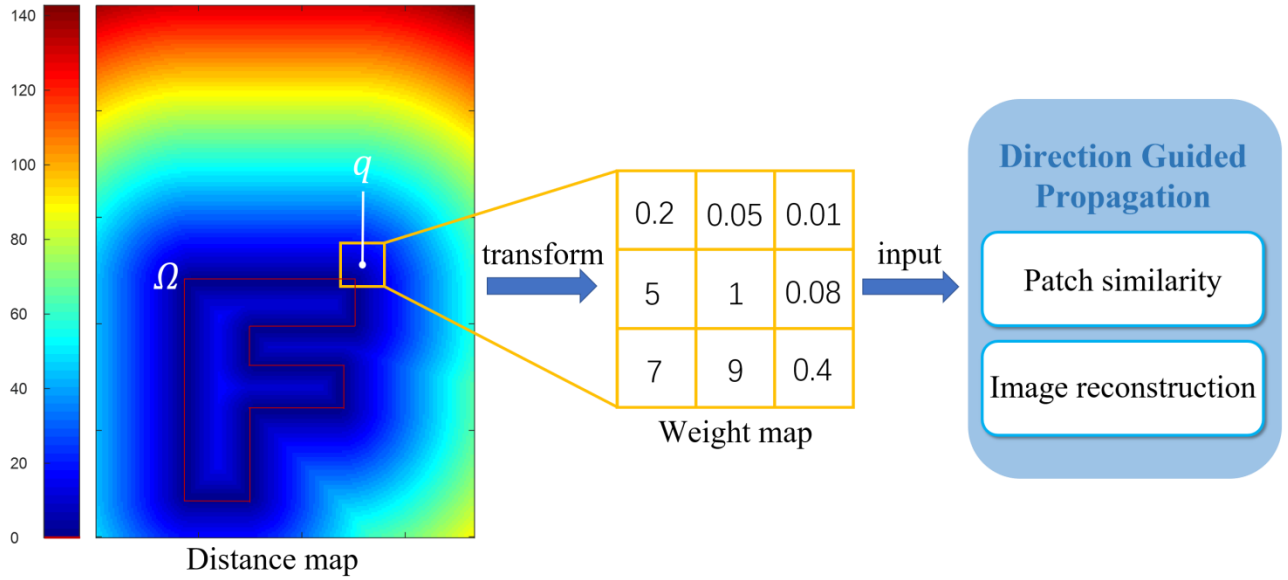


Method

4. Guided Nearest-neighbor Field Search

Direction Guided Propagation

Direct the propagation outward from the text contour for complicated textures without semantic guidance



The weight map based on the distance as

$$\alpha_{q'} = \varphi^{-(d_{\perp}(q',\Omega)-d_{\perp}(q,\Omega))}$$

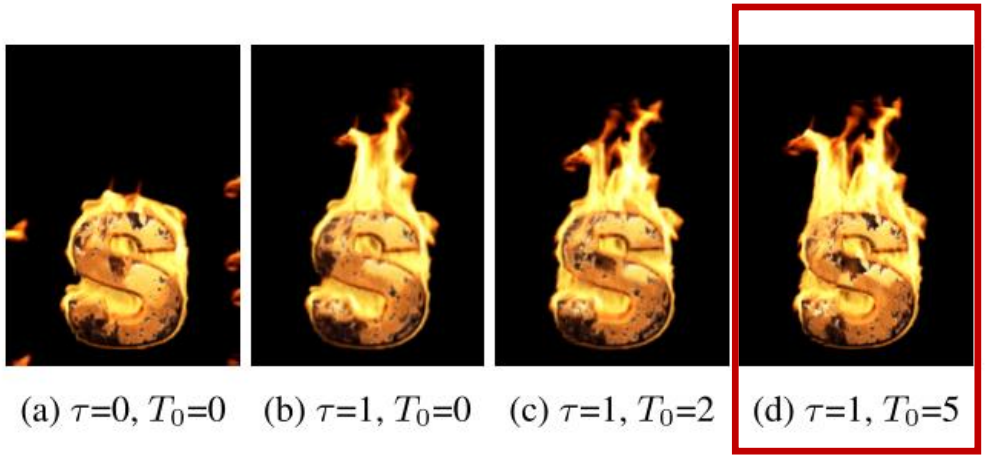
Method

4. Guided Nearest-neighbor Field Search

Simulated Annealing for deep propagation

ALGORITHM 2: Propagation with Simulated Annealing

Input: $S_{text}, T_{text}, S_{sty}, T_{sty}, NN, idirect, \varphi_{cur}, \varphi_{total}, T_0, T_f$
Output: Nearest neighbor fields NN
repeat
  Generate candidate solution NN' and compute its energy value E' using the weight map;
  Compute $\Delta E = E' - E$;
  if $\Delta E < 0$ **then**
    Update nearest neighbor $NN = NN'$;
  else
    Set $\Delta T = \frac{\varphi_{cur}}{\varphi_{total}}(T_0 - T_f)$;
    Compute acceptance probability $prob = min(1, exp\{-\frac{\Delta E}{T_0 - \Delta T}\})$;
    if $prob > \xi(random(0,1))$ **then**
      Update nearest neighbor $NN = NN'$;
    end
  end
until $nUpdate > 0$;



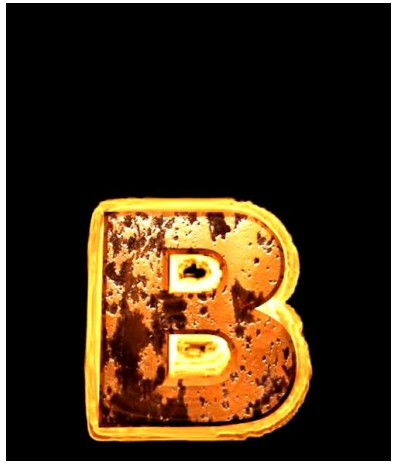
(a) τ=0, T_0=0 (b) τ=1, T_0=0 (c) τ=1, T_0=2 (d) τ=1, T_0=5

Without WM&SA | Without SA | With WM and SA

Results

Example-based Dynamic Text Effects Transfer

target text (input)



source video (input)

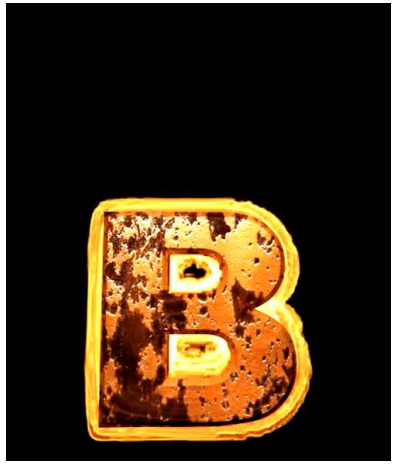
background (input)

target video (output)

Results

Example-based Dynamic Text Effects Transfer

target text (input)



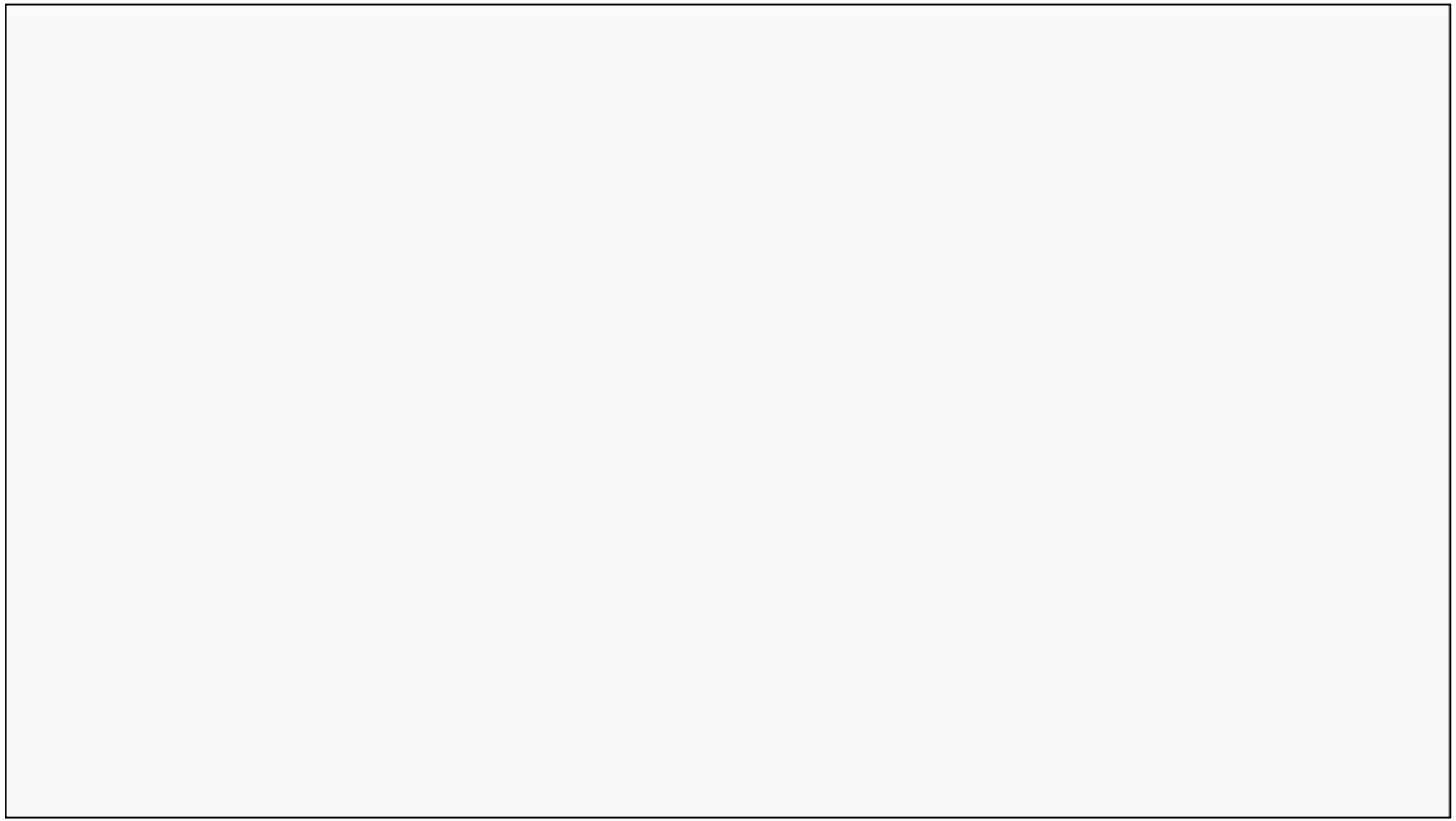
source video (input)

background (input)

target video (output)

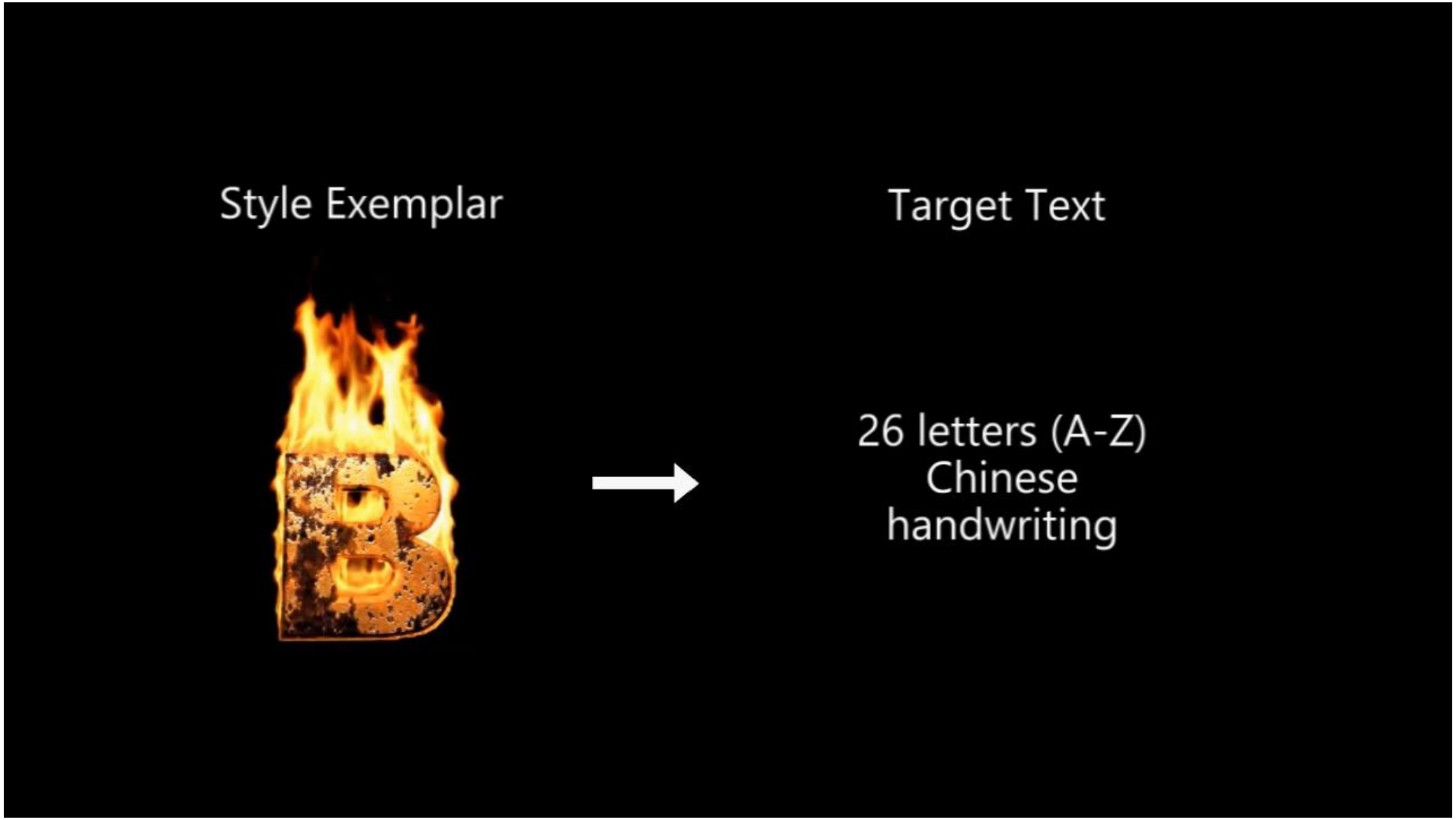
Results

Comparisons



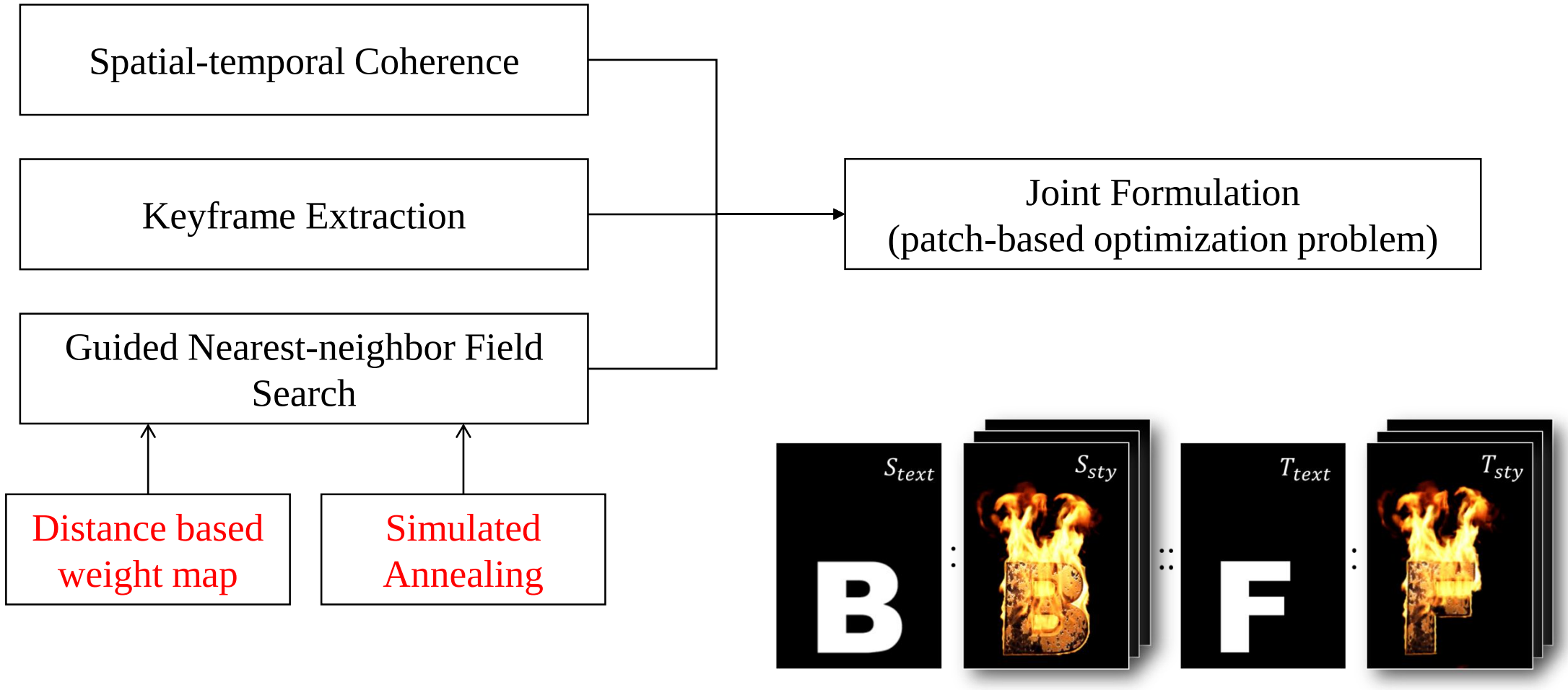
Results

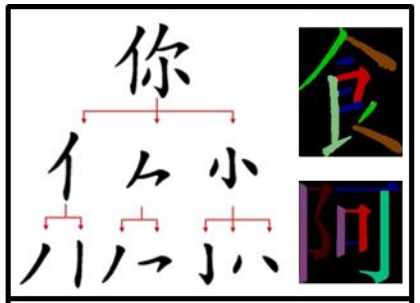
More Results



Conclusion

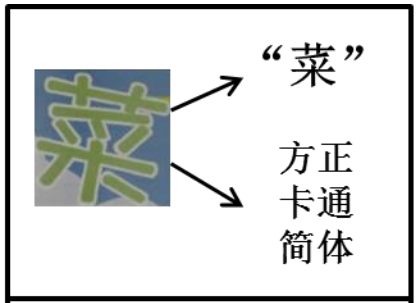
Revisiting





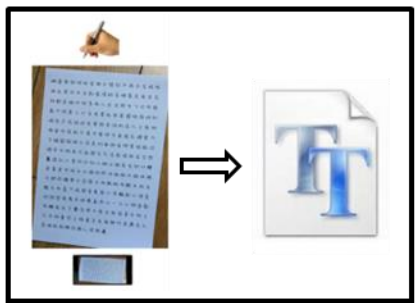
字形切分

文本定位



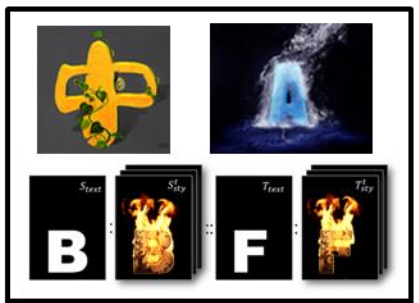
字形识别

识字认字



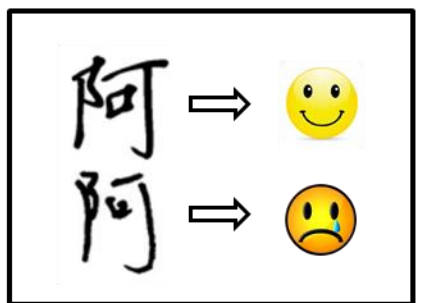
字库生成

书法创作



特效字体

写字做字



字形评价

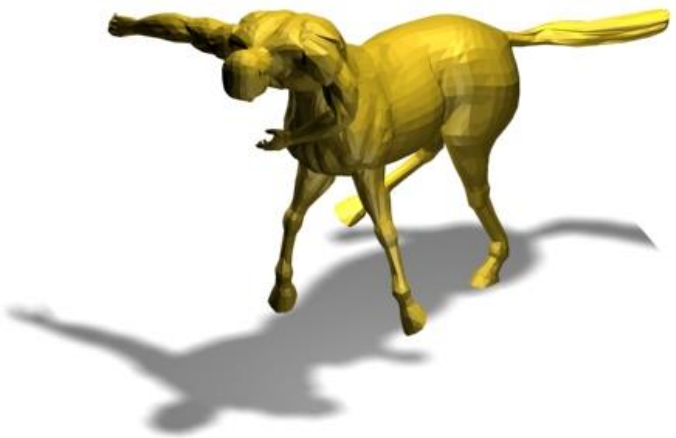
个性社交

平面媒体

赏字用字

北京大学
PEKING UNIVERSITY
计算机科学技术研究所
字形计算技术实验室
http://www.icst.pku.edu.cn/cscl
lianzhouhui@pku.edu.cn
连宙辉

Thank You



北京大学
PEKING UNIVERSITY
计算机科学技术研究所
字形计算技术实验室
http://www.icst.pku.edu.cn/cscl
lianzhouhui@pku.edu.cn
连宙辉