



Методы повышения визуального качества восстановленного фона (часть 2)

Дмитрий АКИМОВ

Video Group
CS MSU Graphics & Media Lab



Содержание

- **Введение**
- Coherent Occlusion Generation
- Background Inpainting
- Gradient Domain Video Editing
- Заключение

Введение

Постановка задачи

- Дано:
 - Исходное видео
 - Карта глубины
- Требуется:
 - Определить области открытия, появляющиеся при генерации стерео
 - Заполнить найденные области таким образом, чтобы человек не был способен обнаружить следов редактирования

Введение

Текущие результаты (1)

Source





Введение

Текущие результаты (1)

Source



Кадр №78



Введение

Текущие результаты (1)

Depth



Кадр №78

Введение

Текущие результаты (1)

Occlusion mask



Кадр №78



Введение

Текущие результаты (1)

Reconstruction result



Кадр №78

Введение

Текущие результаты (1)

Source



Кадр №78

Введение

Текущие результаты (2)

Source



Введение

Текущие результаты (2)

Source



Кадр №7



Введение

Текущие результаты (2)

Depth



Кадр №7

Введение

Текущие результаты (2)

Occlusion mask



Кадр №7

Введение

Текущие результаты (2)

Reconstruction result



Кадр №7

Введение

Текущие результаты (2)

Source



Кадр №7

Введение

Текущие результаты (3)



Текущие результаты

Проблемы

- Неточность результатов Motion Estimation
- Некорректная обработка движущихся объектов заднего плана
- Нехватка информации для восстановления в выбранном временном окне



Содержание

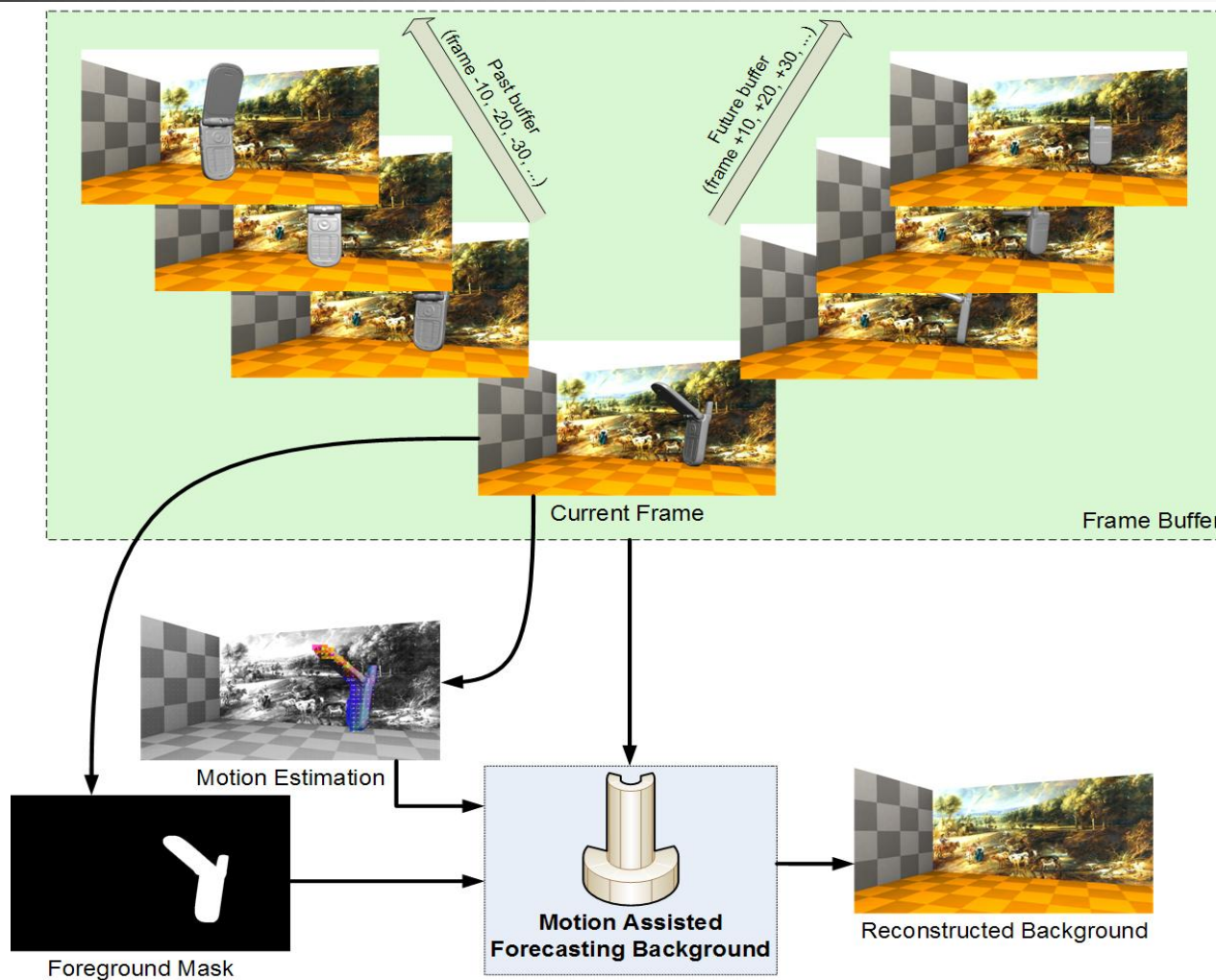
- Введение
- **Coherent Occlusion Generation**
- Background Inpainting
- Gradient Domain Video Editing
- Заключение

Coherent Occlusion Generation



- Дано:
 - Исходное видео
 - Карта глубины
- Требуется восстановить стабильные во времени фон и карту глубины фона в областях открытия для последующей генерации многокурсного видео

Краткая схема метода



Краткая схема метода

Motion Assisted Forecasting Background

1. Вычисление векторов движения между текущим кадром и кадрами из временной окрестности
2. Для каждого пикселя неизвестной области:
 - 1) Поиск векторов-кандидатов из пространственной окрестности
 - 2) Вычисление весов пикселей-кандидатов, соответствующих отобранному вектору
 - 3) Взвешенное суммирование пикселей-кандидатов

Motion Assisted Forecasting

Поиск векторов-кандидатов



В окрестности пикселя $p(x, y, t_0)$ ищется первый вектор движения (v_x, v_y, t) фона, для которого пиксель $c(x + v_x, y + v_y, t_0 + t)$ находится на фоне и не в области закрытия

Может быть найден один кандидат из предыдущих кадров и один из последующих кадров

Motion Assisted Forecasting

Результат заполнения



Если для p найдены кандидаты, то результат заполнения r вычисляется по формулам:

$$r(x, y, t_0) = \sum_{c \in Cand} \sum_{\substack{\omega(u, v) \in \\ \sigma_{(s_x, s_y)}(c(x+u_x, y+v_y, t_0+t))}} \omega(u, v) \cdot 2^{-10^{1-\text{mixFactor}} \cdot (\text{worst}-\text{best}) \cdot (\omega(u, v) - \text{best})}$$

$$\omega(u, v) = \sum_{w \in W_x} \sum_{v' \in W_y} \sum_{\tau \in W_{t_0}} \sum_{k \in \{R, G, B, Depth\}} W(p(x + u', y + v', \tau, k), c(u' + u_x, v' + v_y, \tau + t,))$$

$$W(a, b) = \begin{cases} |a - b| & \text{if } a, b \in \text{background} \\ \text{penalty} & \text{otherwise} \end{cases}$$

$Cand$ – множество пикселей-кандидатов

mixFactor , penalty – константные параметры

worst , best – максимальное и минимальное значение ω для данного p

W_x , W_y , W_{t_0} – размеры пространственно-временной окрестности пикселя

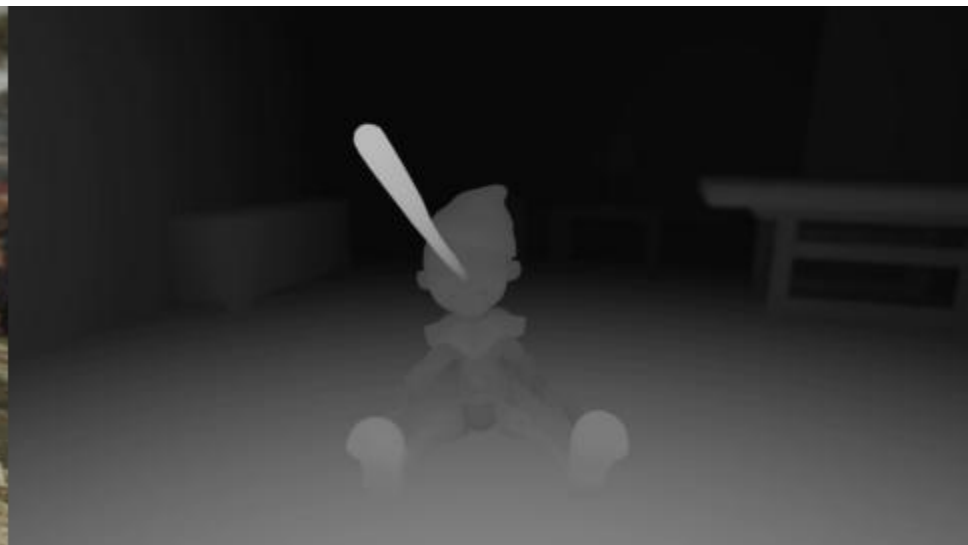
$\sigma_{(s_x, s_y)}(c)$ – область вокруг пикселя-кандидата

Occlusion Generation

Результаты (1)



Исходный кадр



Карта глубины

Occlusion Generation

Результаты (1)



Восстановленный фон



Карта глубины фона

Occlusion Generation

Результаты (2)



Исходный кадр

Occlusion Generation

Результаты (2)



Восстановленный фон

Occlusion Generation

Выводы

- **Достоинства:**
 - Straight forward подход
 - Стабильность во времени (по заявлениям авторов)
- **Недостатки:**
 - Результаты только на искусственных последовательностях
 - Размытие фона (вытекает из усреднения по окрестности найденного кандидата)



Содержание

- Введение
- Coherent Occlusion Generation
- **Background Inpainting**
- Gradient Domain Video Editing
- Заключение

Background Inpainting

Постановка задачи

- Дано:
 - Исходное видео (статичный фон, камера и объекты двигаются произвольно)
 - Маска объектов, которые требуется удалить
 - Маска объектов, которые требуется оставить
- На выход:
 - Отретушированное видео без «ненужных» объектов

Background Inpainting

Краткий обзор

Background Inpainting for Videos with Dynamic Objects and a Free-moving Camera

M. Granados¹ K. I. Kim¹ J. Tompkin^{1,2,3}
J. Kautz² C. Theobalt¹

¹MPI Informatik

²UCL

³Intel Visual Computing Institute

Background Inpainting

Сопоставление кадров

Кусочно-линейное проективное преобразование целевого кадра V_t в заполняемый V_s :

1. Сопоставление feature points (использовались KLT (Kanade-Lucas-Tomasi Feature Tracker) [3,4] и SURF (Seeded-Up Robust Features) [5])
2. Построение множества $H_{st} = \{H_{st}^1, \dots, H_{st}^k\}$ проективных преобразований
3. Вычисление $K(p)$ – функции сопоставления заполняемой точке p оптимального преобразования H_{st}^k

Background Inpainting

Множество преобразований

■ Для каждой пары кадров на основе найденных feature point вычисляется набор проективных преобразований:

- Шаг 1 – учет всех точек, вычисление преобразования и outlier feature points алгоритмом RANSAC
- ...
- Шаг n – учет outlier точек шага $n - 1$, вычисление преобразования

RANSAC

Краткий экскурс

Дано:

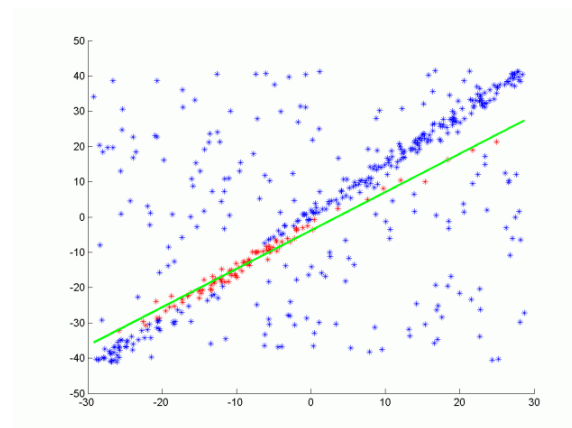
- Набор данных X
- Модель данных $M(\theta)$
- Функция $\theta(X')$ задания параметров модели на основе набора данных
- Функция $E(x, \theta)$ соответствия x модели $M(\theta)$
- Параметры t, k – порог для функции E и числа итераций алгоритма

Результат:

- Оптимальные параметры θ

Алгоритм для i -го шага:

1. Случайно выбрать подмножество $X' \subset X$
2. Вычислить $\theta_i = \max(\theta_i(X'), \theta_{i-1})$
3. Обновить метки множества X



Пример работы алгоритма для поиска прямой на плоскости

Background Inpainting

Финальное преобразование (1)

- Итоговое преобразование F_{st} задается множеством H_{st} и функцией $K(p)$
- $K(p)$ вычисляется методом GraphCut для функционала:

$$\mathcal{E}(K) = \sum_{p \in \Omega} E_p^1(K(p)) + \beta \sum_{(p,q) \in \mathcal{N}(\Omega)} E_{p,q}^2(K(p), K(q))$$

E_p^1 – ограничение на качество преобразования данного пикселя

$E_{p,q}^2$ – ограничение на гладкость

Ω – область изображения

β – константный параметр

(влияет на соотношение связности/качества решения)

$(p,q) \in \mathcal{N}(\cdot)$ – индикатор того, что p и q соседи (по 4-связной системе)

Background Inpainting

Финальное преобразование (2)

Веса функционала:

$$E_p^1(k) = C_{st}^k(p) \cdot \|V_t(p) - V_s(H_{st}^k p_h)\|_2$$

$$E_{p,q}^2(u, v) = \mathbf{1}_{\{u \neq v\}} (\|V_s(H_{st}^u p_h) - V_s(H_{st}^v p_h)\|_2 + \|V_s(H_{st}^u q_h) - V_s(H_{st}^v q_h)\|_2)$$

$$C_{st}^k(p) = 1 - \left[\exp \left(-\frac{1}{2} \frac{(p_h f_{st} H_{st}^k p_h)^2}{r^2} \right) - \frac{1}{2} \right]$$

p_h – точка в однородной системе координат

f_{st} – фундаментальная матрица преобразования кадров

r – константный параметр

Background Inpainting

Заполнение областей (1)

- Для каждого пикселя p заполняемой области функция S_t определяет кадр, из которого будет произведено заполнение
- $S_t(p)$ вычисляется методом GraphCut для функционала:

$$\mathcal{E}'(S_t) = \sum_{p \in \mathcal{H}_t} E'_p{}^1(S_t(p)) + \gamma \sum_{(p,q) \in \mathcal{N}(\mathcal{H}_t \cup \partial \mathcal{H}_t)} E'_{p,q}{}^2(S_t(p), S_t(q))$$

$E'_p{}^1$ – ограничение на качество преобразования

$E'_{p,q}{}^2$ – ограничение на гладкость

$\mathcal{H}_t \cup \partial \mathcal{H}_t$ – область ретуши и пограничные пиксели области

γ – константный параметр

$(p,q) \in \mathcal{N}(\cdot)$ – индикатор того, что p и q соседи (по 4-связной системе)

Background Inpainting

Заполнение областей (2)

Веса функционала:

$$E'_p{}^1(u) = \|W_u^t(p) - R_t(p)\|_2$$

$$E'_{p,q}{}^2(u, v) = \mathbf{1}_{\{u \neq v\}} \left(\|W_u^t(p) - W_v^t(p)\|_2 + \|W_u^t(q) - W_v^t(q)\|_2 \right)$$

$$R_t(p) = \frac{\sum_{l=1}^T a_l^t W_l^t(p)}{\sum_{u=1}^T a_l^t}, \quad a_l^t = \exp\left(-\frac{A_l^t}{\sigma_{A^t}}\right), \quad A_l^t = \frac{\sum_{p \in \Omega \setminus \mathcal{H}_t} d_{\mathcal{H}_t}(p) e_l^t(p)}{\sum_{p \in \Omega \setminus \mathcal{H}_t} d_{\mathcal{H}_t}(p)}$$

$$d_{\mathcal{H}_t}(p) = \exp\left(-\frac{1}{2} \frac{D(p, \mathcal{H}_t)}{\sigma_d}\right)$$

$$W_u^t \equiv V_u(F_{ut})$$

R_t – заполнение области взвешенным средним по всем кадрам

σ_{A^t} – дисперсия $\{A_l^t\}_{l=1..T}$

D – distance transform преобразование области

Background Inpainting

Финальная ретушь

С помощью метода Poisson Image Editing подавляются швы заполнения и различия в освещенности

Функционал:

$$\min_f \sum_{(p,q) \in \mathcal{N}(\Omega)} [(f_p - f_q) - v_{pq}]^2 + \lambda \sum_{p \in \Omega} (f_p - g_p^*)^2$$

Минимум находится из СЛАУ:

$$(|\mathcal{N}_p| + \lambda) f_p - \sum_{q \in \mathcal{N}_p^{\mathcal{H}}} f_q = \sum_{q \in \mathcal{N}_p^{\mathcal{H}}} v_{pq} + \sum_{q \in \mathcal{N}_p^{\partial \mathcal{H}}} f_q^* + \lambda g_p^*$$

Background Inpainting

Результаты



Results

Background Inpainting

Выводы

- **Достоинства:**
 - Качественные результаты
 - Стабильность во времени
 - Применим для некалиброванной камеры
- **Недостатки:**
 - Восстанавливает статичный фон
 - Требуется качественная разметка данных



Содержание

- Введение
- Coherent Occlusion Generation
- Background Inpainting
- **Gradient Domain Video Editing**
- Заключение

Gradient Domain Editing

Poisson Image Editing

Задача ретуши изображения:

$$\min_u \int_{O \subset \Omega} \|\nabla u - g\|^2 dx; \quad \text{with} \quad u|_{\partial O} = u_0$$

u – исходное изображение

g – градиент вставляемого изображения

Ω – пространство изображения

O – область ретуширования

Решение ищется в виде вариационной задачи для уравнения Пуассона с граничными условиями Дирихле

Poisson Image Editing

Результат



Исходные изображения



Заполнение



Результат

Gradient Domain Video Editing

Переформулировка задачи для случая видео:

$$E_p(u) = \int_{O \subset \Pi} |\partial_v u(x, t)|^p + \beta \|\nabla_x u(x, t) - g(x, t)\|^p dx dt$$

u – исходное видео

g – градиент вставляемого видео

v – оптический поток

∂_v – производная вдоль
направления потока

$\Pi = \Omega \times T$ – пространственно-временной
домен видео

O – область ретуширования

β, p – параметры модели

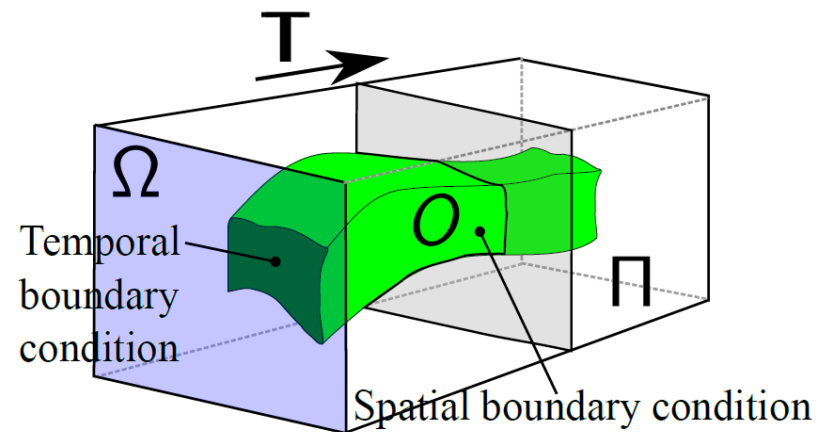


Иллюстрация пространственно-
временного домена

Вариационная задача

Производная по направлению

■ Приближения производной по направлению с помощью forward v^+ и backward v^- OF [7]:

$$\partial_v^+ u(x, t) := \begin{cases} \hat{u}(x + v^+(x, t), t + 1) - u(x, t) & \text{if } t < T, \\ 0 & \text{if } t = T \text{ or } (x + v^+(x, t), t + 1) \notin O, \end{cases}$$

$$\partial_v^- u(x, t) := \begin{cases} \hat{u}(x, t) - u(x + v^-(x, t), t - 1) & \text{if } t > 0, \\ 0 & \text{if } t = T \text{ or } (x + v^-(x, t), t - 1) \notin O. \end{cases}$$

$$Occ(x, t) = \begin{cases} 1 & \text{if } |\hat{u}_0(x + v^+(x, t), t + 1) - u_0(x, t)| > \text{const} \\ 0 & \text{otherwise} \end{cases}$$

$$\bar{\partial}_v \in \{\partial_v^+, \partial_v^-\}$$

\hat{u} – билинейная интерполяция u

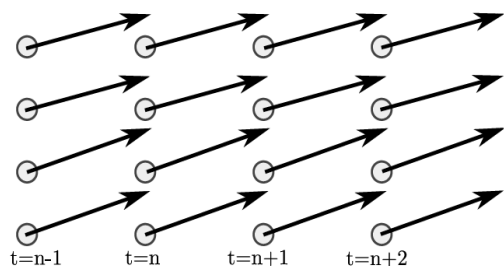
Минимизируемый функционал:

$$E_p(u) = \sum_{t \in \mathbb{T}} \sum_{x \in \Omega} \|Occ(x, t) \bar{\partial}_v u(x, t)\|^p + \beta \sum_{t \in \mathbb{T}} \sum_{x \in \Omega} \|\nabla_x^+ u(x, t) - g(x, t)\|^p$$

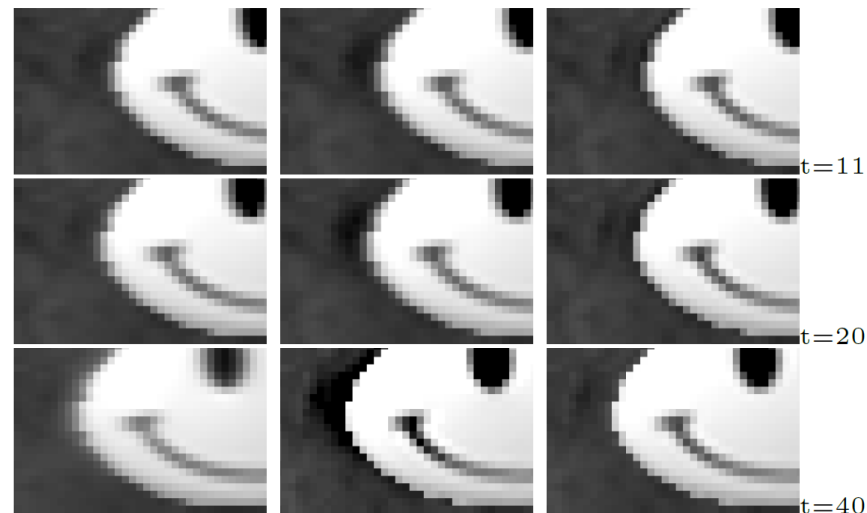
Вариационная задача

Deblurring Convective Derivative (DCD)

Использование только v^+ приближения приводит к осцилляции u в домене Π , только v^- – к сглаживанию. Предлагается схема чередования приближений:



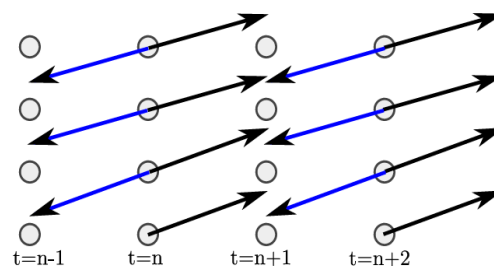
v^+ -scheme



v^- -scheme

v^+ -scheme

DCD



DCD

Вариационная задача

DCD – сравнение

V+ scheme



V- scheme



DCD scheme



Gradient Domain Editing

Результаты (1)

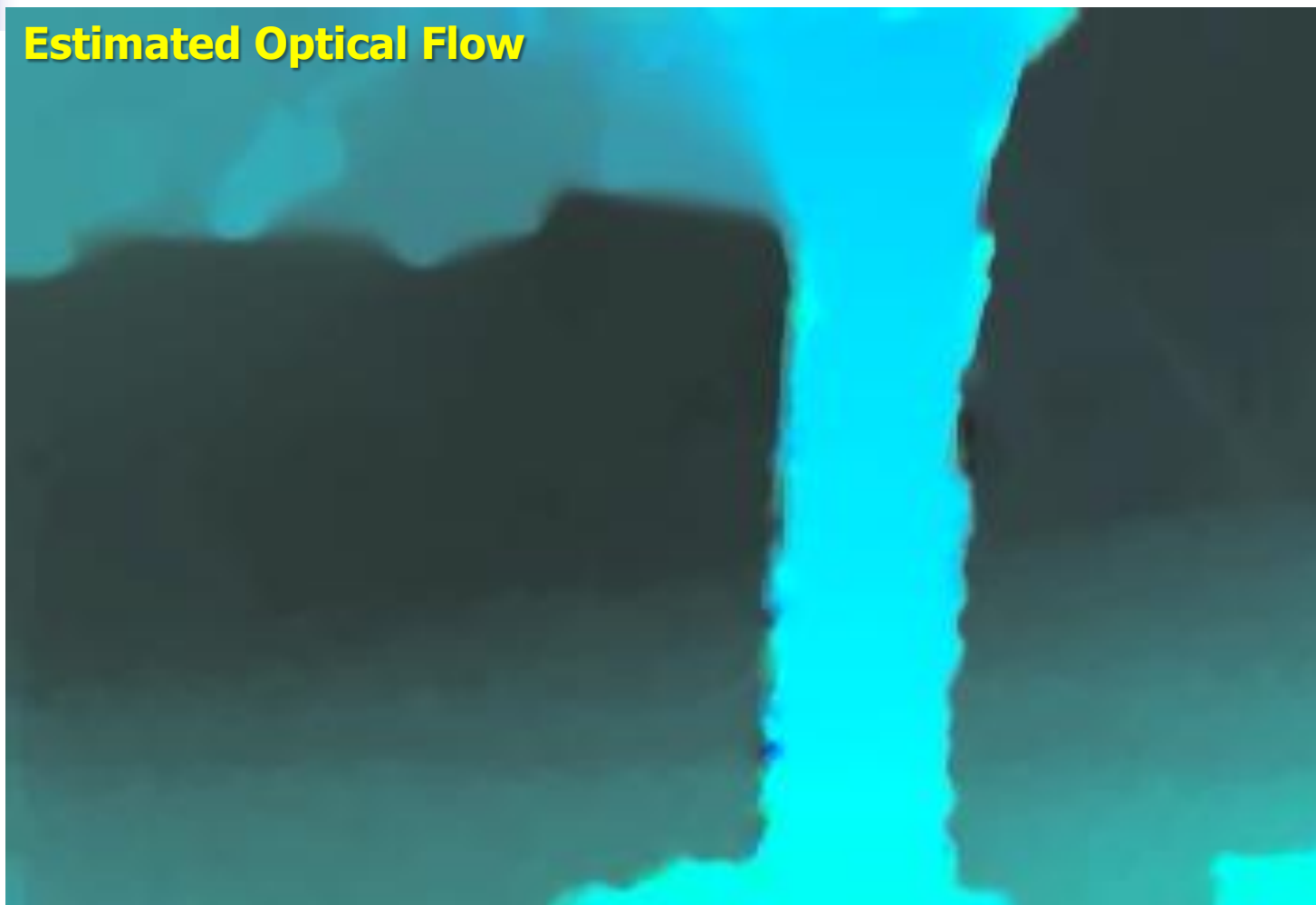
Source



Gradient Domain Editing

Результаты (1)

Estimated Optical Flow



Gradient Domain Editing

Результаты (1)

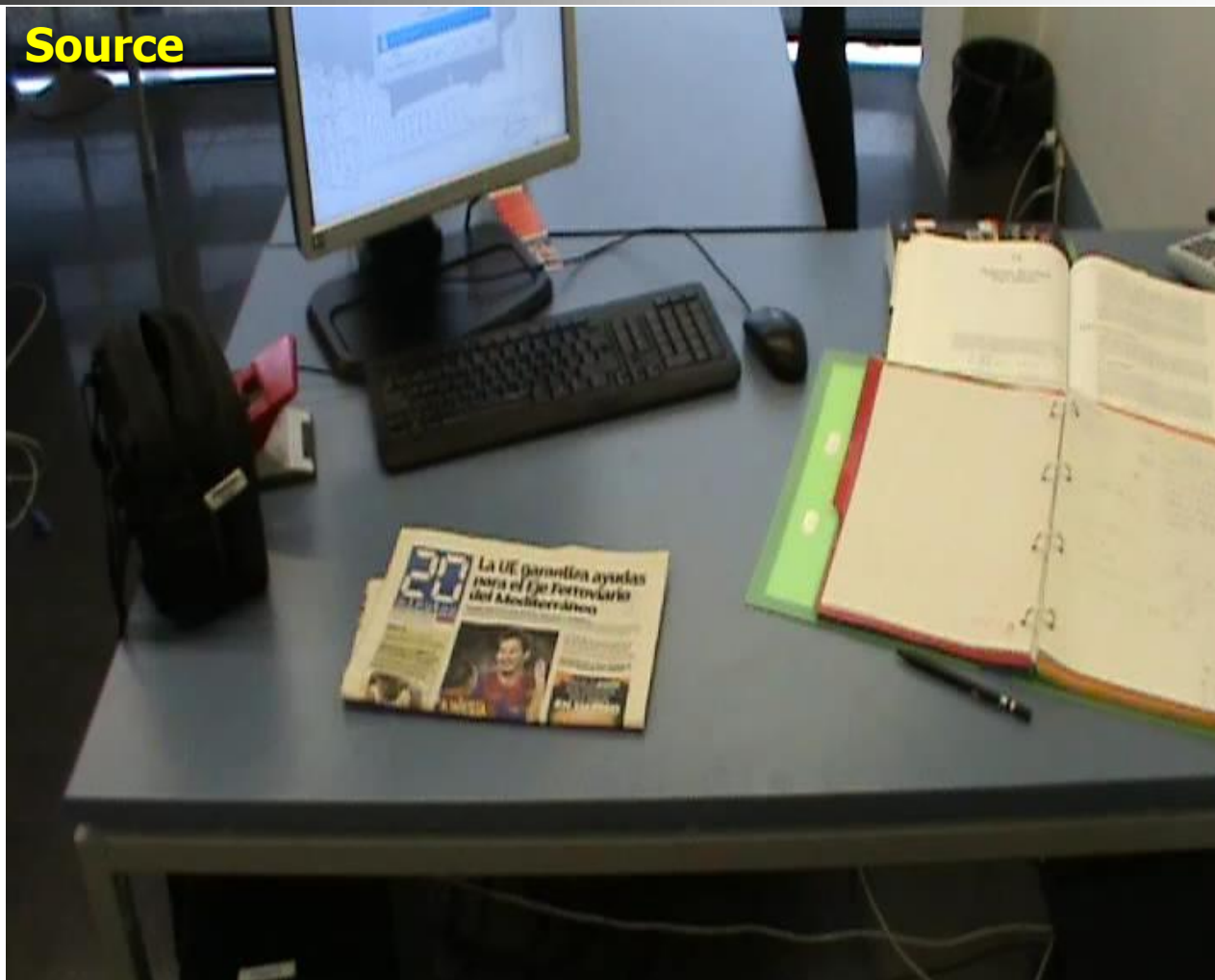
Result



Gradient Domain Editing

Результаты (1)

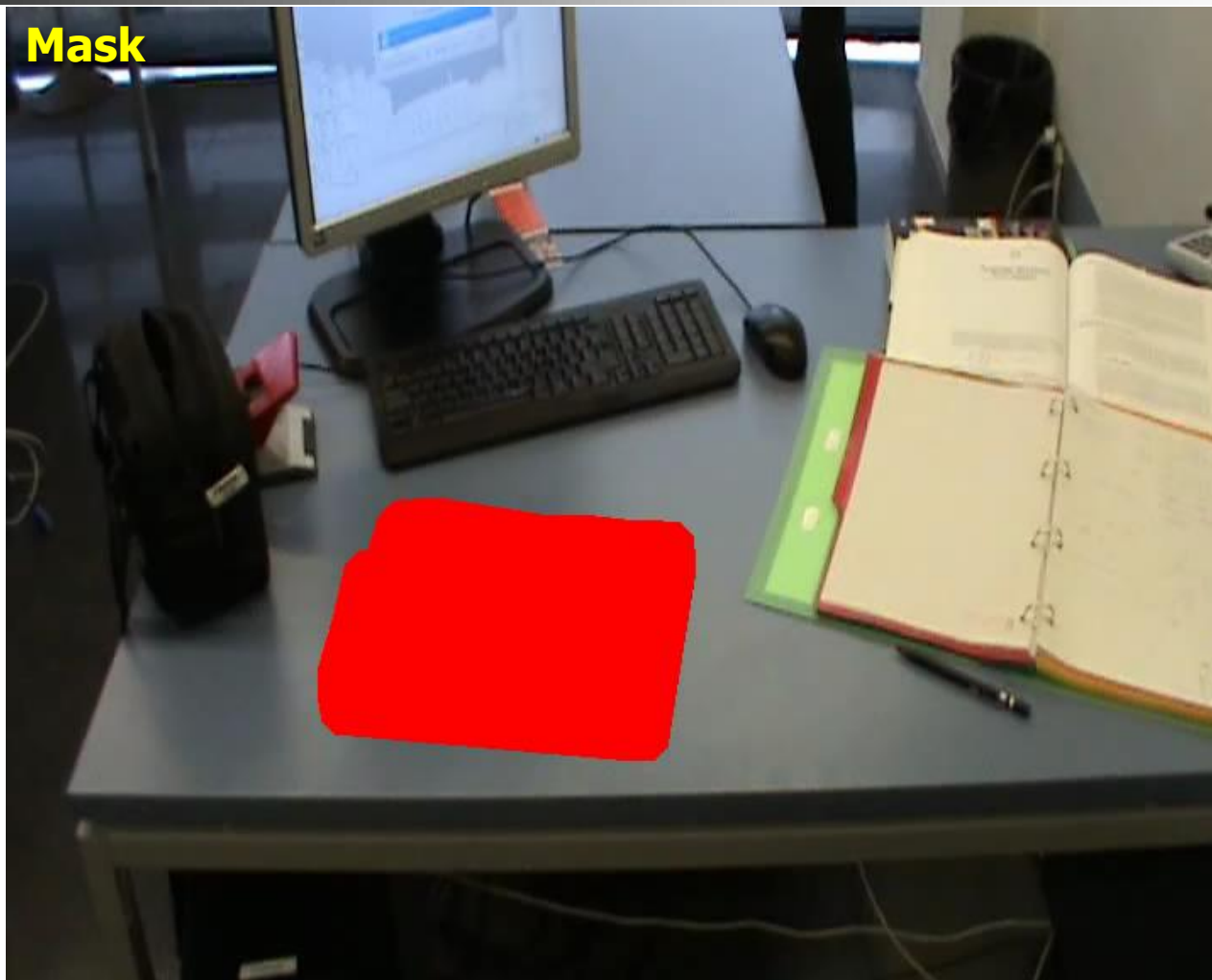
Source



Gradient Domain Editing

Результаты (1)

Mask



Gradient Domain Editing

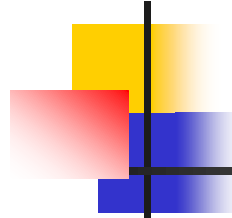
Результаты (1)

Result



Gradient Domain Editing

Результаты (2)

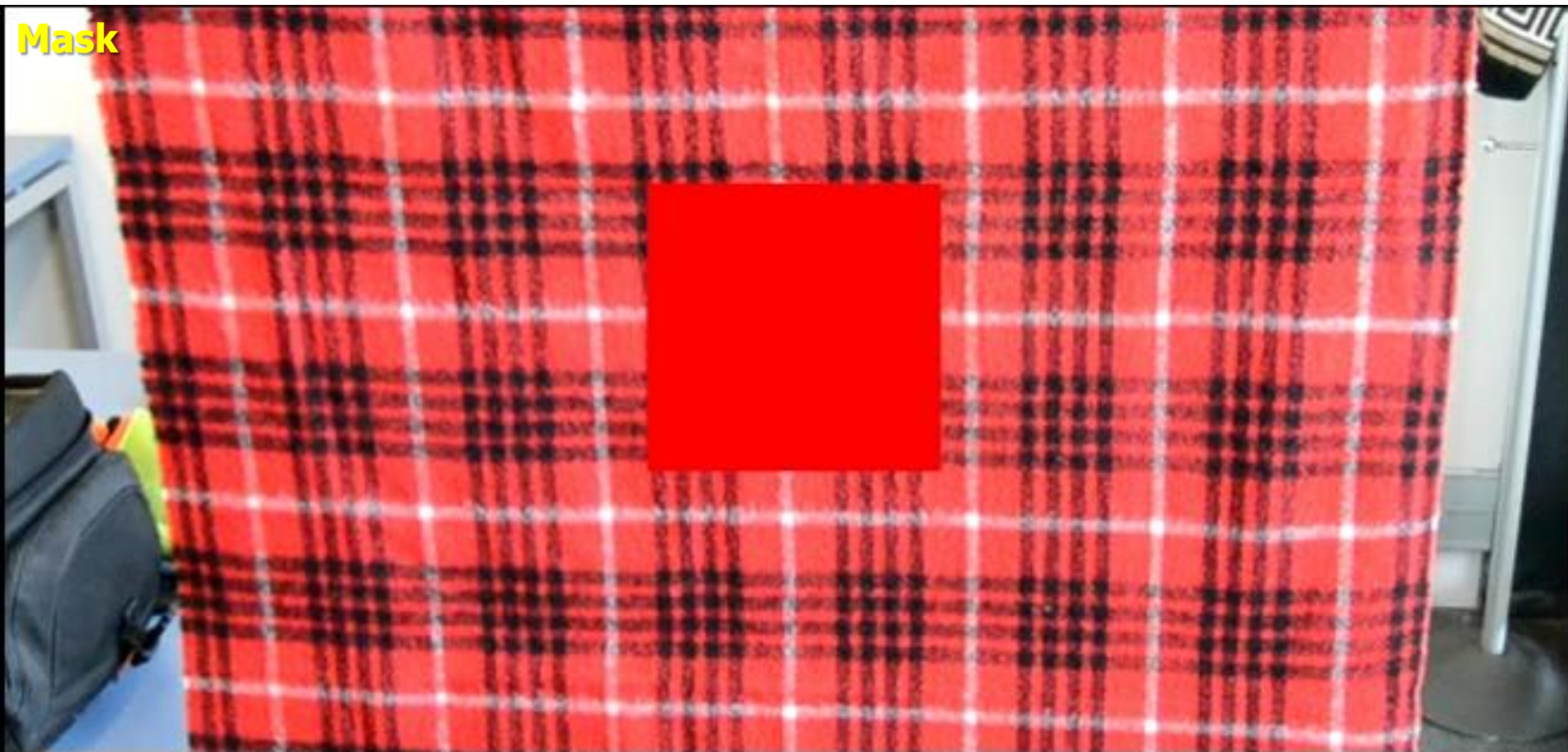


Source



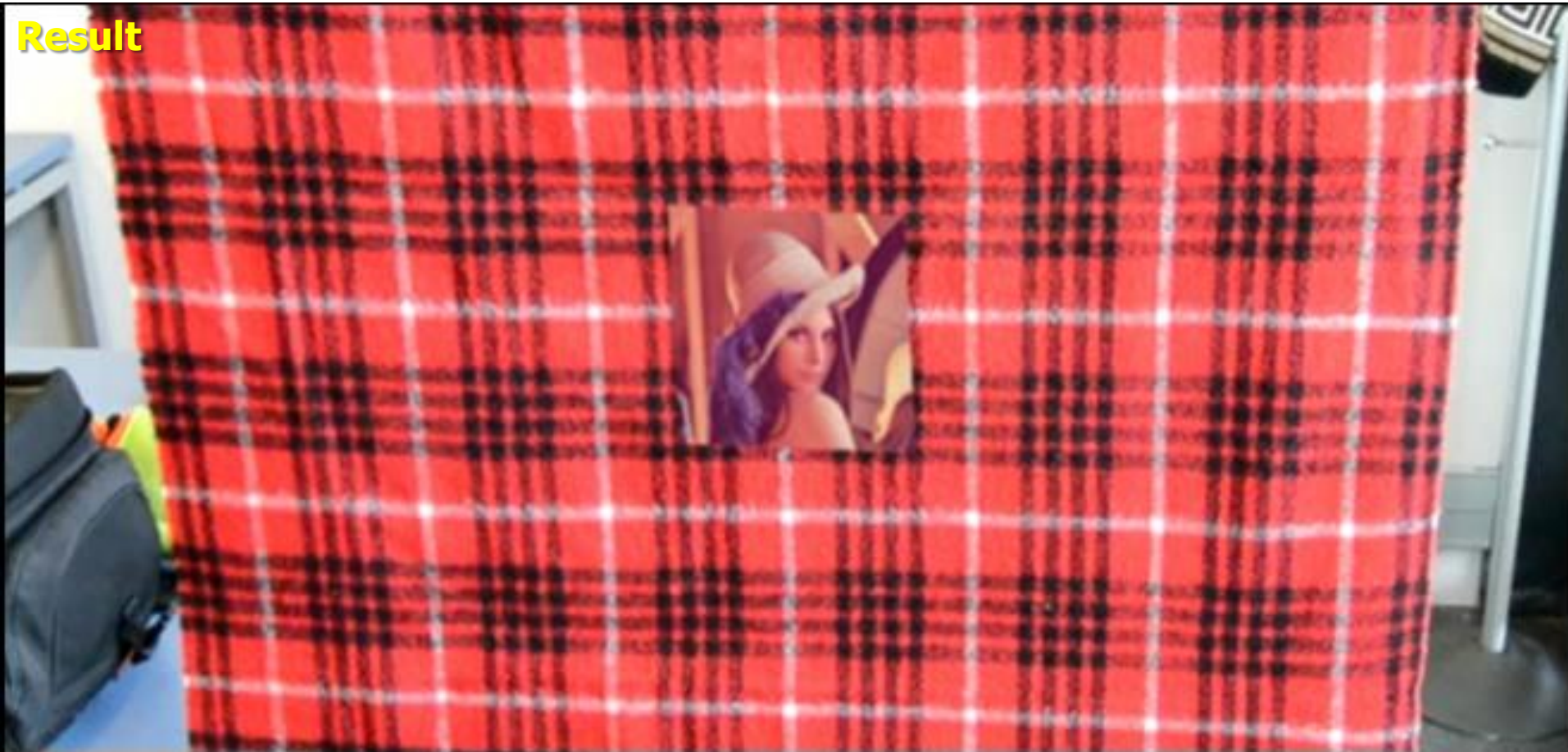
Gradient Domain Editing

Результаты (2)



Gradient Domain Editing

Результаты (2)

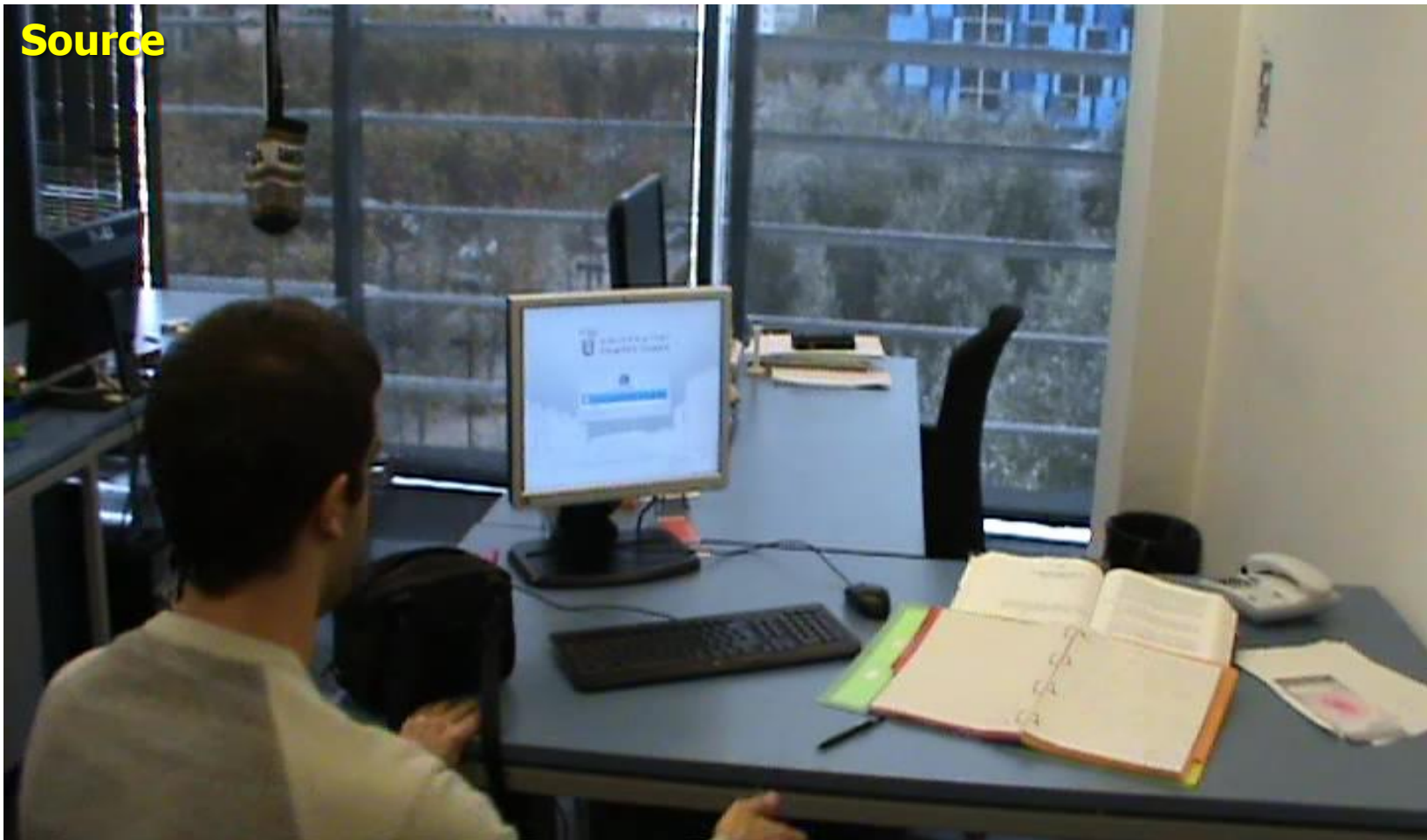


Gradient Domain Editing

Результаты (3)

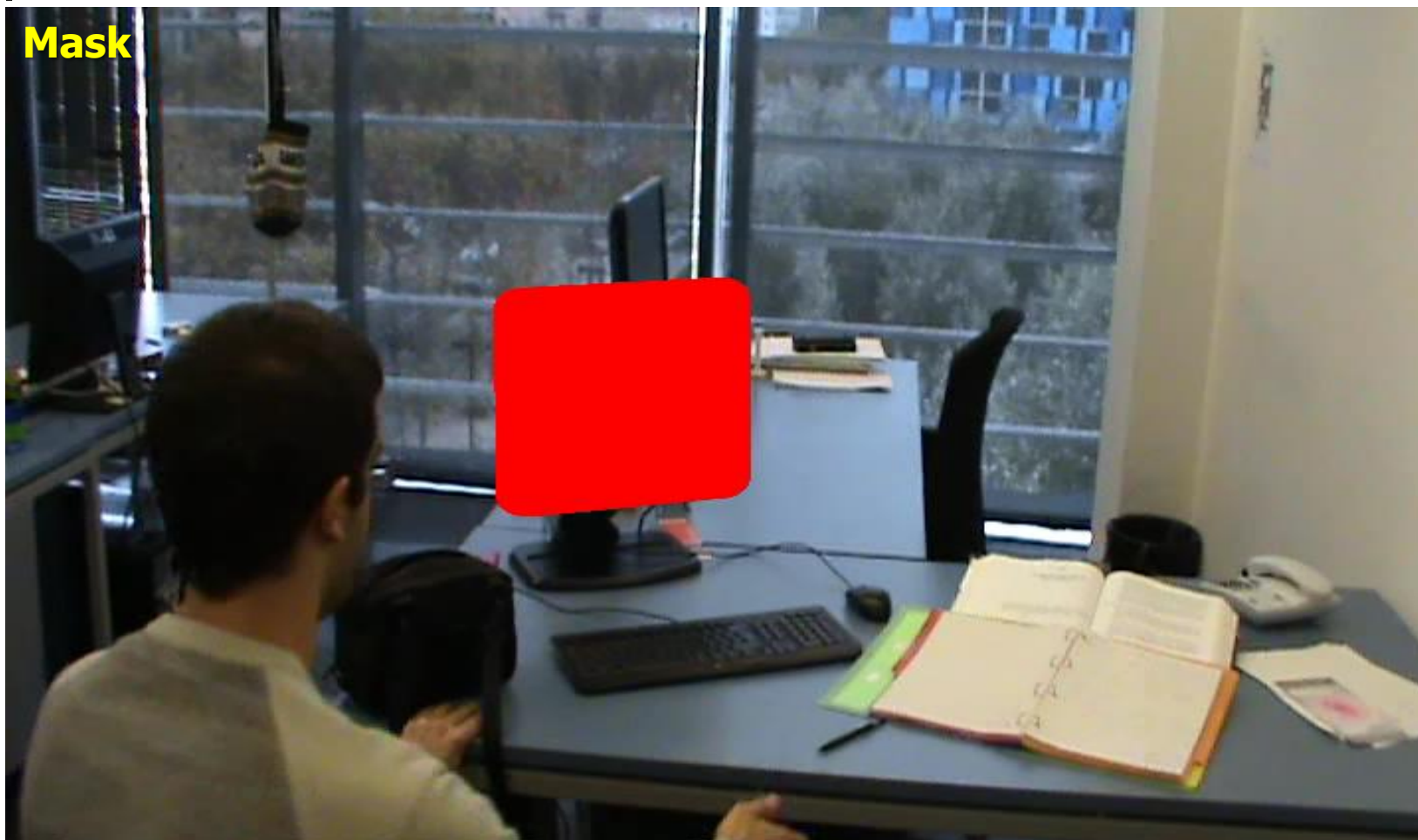


Source



Gradient Domain Editing

Результаты (3)



Gradient Domain Editing

Результаты (3)



Result



Gradient-Domain Editing

Выводы

- **Достоинства:**
 - Качественные результаты
 - Стабильность во времени
 - Применим для некалиброванной камеры
- **Недостатки:**
 - Высокая вычислительная сложность
 - Требуется качественная оценка движения

Заключение

- Рассматриваемая задача критична в задачах качественной генерации и коррекции стерео
- На данный момент нет решения, качественно и напрямую решающего поставленную задачу
- Наиболее интересны подходы:
 - Кусочно-линейное проецирование кандидатов
 - Обработка видео в градиентном домене

Литература

1. R. K. Gunnewiek, R.-P. M. Berretty, B. Barenbrug, and J. P. Magalhães, "Coherent spatial and temporal occlusion generation," in *Proceedings of SPIE*, 2009.
2. M. Granados, K. L. Kim, J. Tompkin, J. Kautz, C. Theobalt, "Background Inpainting for Videos with Dynamic Objects and a Free-Moving Camera," in *Proceedings of ECCV*, 2012, pp. 682-695.
3. B.D. Lucas, T. Kanade, "An Iterative Image Registration Technique with an Application to Stereo Vision," in *Proceeding of IJCAI*, 1981, pp. 674-679.
4. J, Shi, C. Tomasi, "Good Features to Track," in *Proceedings of IEEE CVPR*, 1994, pp. 593-600.

Литература

5. H. Bay, A. Ess, T. Tuytelaars, L.J.V. Gool, "Speeded-Up Robust Features (SURF), in *Proceedings of Computer Vision and Image Understanding*, 2008, 346-359.
6. G. Facciolo, R. Sadek, A. Bugeau, and V. Caselles, "Temporally consistent gradient domain video editing," in *Proceedings of Energy Minimization Methods in Computer Vision and Pattern Recognition*, 2011, pp. 59-73.
7. T. Brox, A. Bruhn, N. Papenberg, J. Weickert, "High Accuracy Optical Flow Estimation Based on a Theory for Warping," in *Proceedings of European Conference on Computer Vision*, 2004.

Лаборатория компьютерной графики и мультимедиа



Видеогруппа — это:

- Выпускники в аспирантурах Англии, Франции, Швейцарии (в России в МГУ и ИПМ им. Келдыша)
- Выпускниками защищены 5 диссертаций
- Наиболее популярные в мире сравнения видеокодеков
- Более 3 миллионов скачанных фильтров обработки видео