

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Fields of Experts

доклад

1 апреля 2011

Outline

Введение

- Image priors
- Модели патчей

Fields of Experts

- Модель
- Вывод
- Обучение

Приложения

- Image denoising
- Image inpainting

Заключение

Введение

- Image priors
- Модели патчей

Fields of Experts

- Модель
- Вывод
- Обучение

Приложения

- Denoising
- Inpainting

Заключение

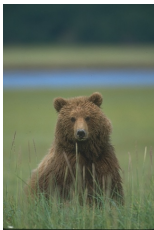
Постановка задачи

Получить модель априорных знаний об изображении
(**image prior**):

$$p(\mathbf{x}), \quad \mathbf{x} \text{ — изображение}$$

содержащую статистику реальных сцен.

Например таких:



Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

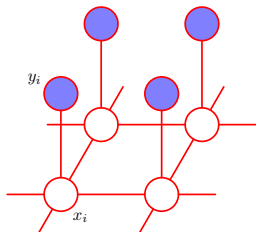
Denoising

Inpainting

Заключение

Где нужны image priors?

- ▶ Задачи низкоуровневого зрения
 - ▶ denoising
 - ▶ inpainting
 - ▶ super-resolution
 - ▶ ...
- ▶ Синтез текстур
- ▶ Коррекция карт глубины
- ▶ Оценка оптического потока
- ▶ Где нибудь еще?



Вероятностные модели структуры изображений обычно формулируются в виде MRF:

$$p(\mathbf{x}) = \frac{1}{Z} \exp \left\{ - \sum_k U_k(\mathbf{x}_{(k)}) \right\}$$

Как правило

- ▶ однородной
- ▶ с парной системой соседства (*pairwise*)
- ▶ в качестве потенциалов — робастная функция разности $U(x_i, x_j) = \rho(x_i - x_j)$

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

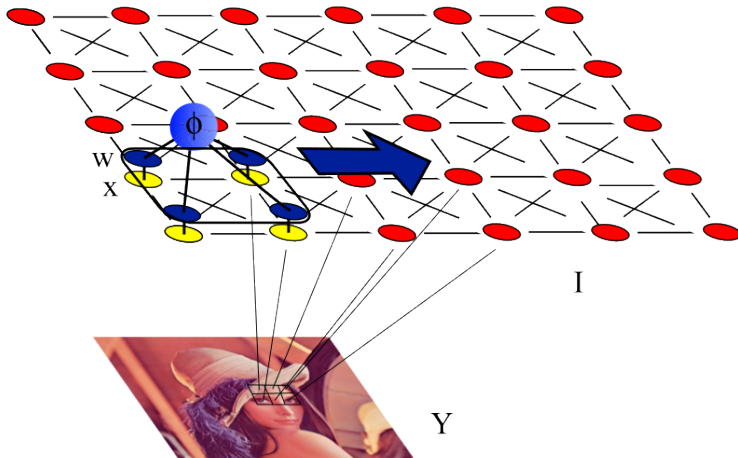
Denoising

Inpainting

Заключение

High-order MRF

Никто не запрещал расширить систему соседства.



Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Вот например, Geman and Reynolds (1992) использовали вторые (и третье) производные на изображении.

Использовались клики размером 3×3 с потенциалами

$$U(\mathbf{x}_{(k)}) = \sum_{i=1}^5 \rho(\mathbf{J}_i^T \mathbf{x}_{(k)}),$$

где \mathbf{J}_i — линейные фильтры производных.

первые производные



вторые производные



Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

- ▶ Ручная настройка параметров модели (фильтров, потенциалов) утомительна и не приводит к желаемому результату.
- ▶ Модель FRAME (Zhu et al., 1998) для генерации текстур использовала high-order MRF. Идея схожа с FoE, однако FRAME для обучения использовала предустановленный набор линейных фильтров.
- ▶ Большинство обучающихся MRF моделей не нашли широкого применения, в основном из-за вычислительных затрат.

Цель авторов FoE: избавиться от ручной настройки чего-либо.

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Небольшие изображения (**патчи**) могут кодироваться откликами линейных фильтров.

$$\mathbf{x} \in \mathbb{R}^n, \quad \mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{J}_i, \quad \alpha_i = \mathbf{J}_i^T \mathbf{x}$$

- ▶ **РСА** позволяет найти ортонормированный базис $\{\mathbf{J}_i\}$, однако компоненты α_i имеют далеко не Гауссово распределение и не независимы.
- ▶ **ІСА** ищет линейные компоненты с минимальной зависимостью, так что можно записать:

$$p(\mathbf{x}) \propto \prod_{i=1}^n p_i(\mathbf{J}_i^T \mathbf{x})$$

- ▶ **Sparse Coding** (Olshausen and Field, 1996)

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Идея PoE: представить многомерное распределение патча в виде произведения «экспертных» распределений, т. ч. каждый эксперт работает в подпространстве меньшей размерности.

$$p(\mathbf{x}; \Theta) = \frac{1}{Z(\Theta)} \prod_{i=1}^n \phi(\mathbf{J}_i^T \mathbf{x}; \alpha_i)$$

$$\Theta = \{\theta_1, \dots, \theta_n\}, \quad \theta_i = (\mathbf{J}_i, \alpha_i)$$

- ▶ Фильтры \mathbf{J}_i обучаются \Rightarrow не нужно беспокоиться о независимости откликов.

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

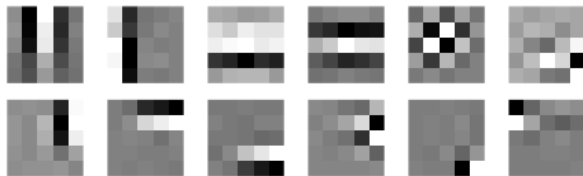
Denoising

Inpainting

Заключение

Products of Experts

Несколько фильтров 5×5 , полученных в результате обучения PoE на базе данных изображений:



Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Outline

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Image denoising

Image inpainting

Заключение

Fields of Experts

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

FoE — однородная MRF высокого порядка с потенциалами PoE на кликах.

$\mathbf{x} \in \mathbb{R}^{L \times M}$; $\mathbf{x}_{(k)}$, $k = 1, \dots, K$ — клики

$$f(\mathbf{x}_{(k)}) = f_{\text{PoE}}(\mathbf{x}_{(k)}; \alpha_i) = \prod_{i=1}^N \phi(\mathbf{J}_i^T \mathbf{x}_{(k)}; \alpha_i)$$

$$p_{\text{FoE}}(\mathbf{x}; \Theta) = \frac{1}{Z(\Theta)} \prod_{k=1}^K \prod_{i=1}^N \phi(\mathbf{J}_i^T \mathbf{x}_{(k)}; \alpha_i)$$

$$\Theta = \{\theta_1, \dots, \theta_n\}, \quad \theta_i = (\mathbf{J}_i, \alpha_i)$$

Θ — обучаемые параметры модели

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

$$\psi(\cdot; \alpha_i) = \log \phi(\cdot; \alpha_i) \quad \textit{log-expert}$$

$$p_{\text{FoE}}(\mathbf{x}; \Theta) = \frac{1}{Z(\Theta)} \exp \{-E_{\text{FoE}}(\mathbf{x}; \Theta)\}$$

$$E_{\text{FoE}}(\mathbf{x}; \Theta) = - \sum_{k=1}^K \sum_{i=1}^N \psi(\mathbf{J}_i^T \mathbf{x}_{(k)}; \alpha_i)$$

- ▶ Гибкость PoE.
- ▶ Моделируются зависимости между перекрывающимися патчами, т.к. параметры Θ обучаются не на независимых патчах, а на изображениях.
- ▶ Инвариантность к сдвигу, т.к. MRF однородна.

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Цветные изображения

- ▶ Преобразовать RGB в YCbCr, использовать каналы отдельно.
- ▶ McAuley et al. (2006) расширили подход авторов FoE для работы с цветными изображениями, используя клики $3 \times 3 \times 3$ и $5 \times 5 \times 3$.
- ▶ Изображения для обучения были преобразованы в YCbCr, использовался только Y-канал.



Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Критерии выбора экспертов $\phi(y; \alpha)$:

- ▶ *Гладкость* по y и α : возможность использования градиентных методов вывода и обучения.
- ▶ *Острота* (kurtotic). Известно, что $\mathbf{J}^T \mathbf{x}$ на реальных изображениях имеет «острое» маргинальное распределение.

Student-t

$$\phi_T(y; \alpha) = \left(1 + \frac{1}{2}y^2\right)^{-\alpha}$$

- ▶ тяжелые хвосты
- ▶ $\psi_T(y)$ НЕ выпуклая

Charbonnier

$$\phi_C(y; \alpha, \beta) = e^{-\alpha\sqrt{\beta+y^2}}$$

- ▶ менее тяжелые хвосты
- ▶ $\psi_T(y)$ выпуклая
- ▶ хорошо сработало в других методах

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

$$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{y}|\mathbf{x})p(\mathbf{x}) \rightarrow \max_{\mathbf{x}}$$

- ▶ У точных методов, а также основанных на ВР нет шансов: сложная структура графа, большие клики.
- ▶ Зато удобно спускаться по градиенту $\log p(\mathbf{x}|\mathbf{y})$:

$$\nabla_{\mathbf{x}} \log p_{\text{FoE}}(\mathbf{x}; \Theta) = \sum_{i=1}^N \mathbf{J}_-^{(i)} * \psi'(\mathbf{J}^{(i)} * \mathbf{x}; \alpha_i) \quad (1)$$

$$\arg \max_{\Theta} p_{\text{FoE}}(X|\Theta), \quad X = \{\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(D)}\}$$

$$\log p(X|\Theta) = - \sum_{i=1}^D E(\mathbf{x}^{(i)}|\Theta) - D \log Z(\Theta)$$

- ▶ Решение в явном виде не выражается
- ▶ Минимизировать log-likelihood не удобно из-за нормировочной константы
- ▶ Зато максимизация правдоподобия эквивалентна минимизации KL

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Минимум KL

Обозначим:

p^0 — распределение данных X ,

p_θ^i — распределение, полученное на i -том шаге MCMC.

p^∞ — истинное распределение $p(x|\theta)$, $p_\theta^n \xrightarrow{n \rightarrow \infty} p_\theta^\infty$

$$\text{KL}(p^0 \| p_\theta^\infty) \rightarrow \min_\theta$$

$$\begin{aligned} \text{KL}(p^0 \| p_\theta^\infty) &= \sum_{\mathbf{x} \in X} p^0(\mathbf{x}) \log p^0(\mathbf{x}) - \sum_{\mathbf{x} \in X} p^0(\mathbf{x}) \log p_\theta^\infty(\mathbf{x}) = \\ &= -H(p^0) - \langle \log p_\theta^\infty \rangle_{p^0} \rightarrow \min_\theta \end{aligned}$$

$$\langle \log p_\theta^\infty \rangle_{p^0} \rightarrow \max_\theta$$

$$\delta\theta_i = \eta \left[\left\langle \frac{\partial E}{\partial \theta_m} \right\rangle_{p_\theta^\infty} - \left\langle \frac{\partial E}{\partial \theta_m} \right\rangle_{p^0} \right] \quad (2)$$

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

$$\text{CD}_n \rightarrow \min_{\theta}$$

$$\text{CD}_n = \text{KL}(p^0 \| p_{\theta}^{\infty}) - \text{KL}(p_{\theta}^n \| p_{\theta}^{\infty})$$

$$\frac{\partial \text{CD}_n}{\partial \theta_m} = \left\langle \frac{\partial E}{\partial \theta_m} \right\rangle_{p^0} - \left\langle \frac{\partial E}{\partial \theta_m} \right\rangle_{p_{\theta}^n} + \frac{\partial p_{\theta}^n}{\partial \theta_m} \frac{\partial \text{KL}(p_{\theta}^n \| p_{\theta}^{\infty})}{\partial p_{\theta}^n} \quad (3)$$

$$\delta \theta_i = \eta \left[\left\langle \frac{\partial E}{\partial \theta_m} \right\rangle_{p_{\theta}^n} - \left\langle \frac{\partial E}{\partial \theta_m} \right\rangle_{p^0} \right]$$

- ▶ На каждой итерации градиентного спуска производится n итераций MCMC

- ▶ Авторы обучали модель с кликами 3×3 и 5×5 на изображениях 15×15
- ▶ Обучающая выборка: 20000 патчей, случайно выбранных из изображений базы Berkley Segmentation Benchmark
- ▶ Выборка разбита на *пачки* по 200 изображений, на каждой итерации используются данные только из одной, случайно выбранной пачки (*стохастический градиентный спуск*)
- ▶ 5000 итераций спуска CD_1 , гибридный метод Монте-Карло.
- ▶ Занимает много времени: Intel Pentium D, 3.2 GHz
 - 3×3 : 8 часов
 - 5×5 : 24 часа

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Outline

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Image denoising

Image inpainting

Заключение

Fields of Experts

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заклучение

$$p(\mathbf{x}) = p_{\text{FoE}}(\mathbf{x})$$

$$\mathbf{y} = \mathbf{x} + \varepsilon, \quad \varepsilon_i \sim \mathcal{N}(0, \sigma^2)$$

$$p(\mathbf{y}|\mathbf{x}) \propto \prod_{k=1}^{L \cdot M} \exp\left(-\frac{1}{2\sigma^2}(y_k - x_k)^2\right)$$

$$\nabla_{\mathbf{x}} \log p(\mathbf{x}|\mathbf{y}) = \omega \cdot \left[\sum_{i=1}^N \mathbf{J}_-^{(i)} * \psi'(\mathbf{J}^{(i)} * \mathbf{x}; \alpha_i) \right] + \frac{1}{\sigma^2}(\mathbf{y} - \mathbf{x})$$

- ▶ $\omega > 0$ регулирует «силу» приора по сравнению с правдоподобием, может быть обучено по проверочной выборке. В реализации авторов, выбирается по σ .

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

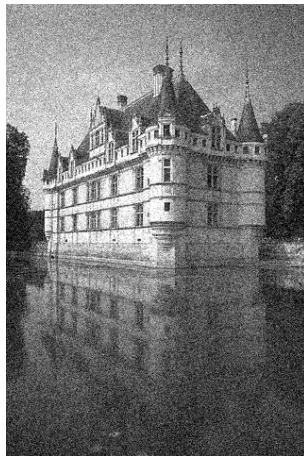
Inpainting

Заключение

Image denoising: сравнение результатов



Оригинал



Зашумленное ($\sigma = 25$)

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

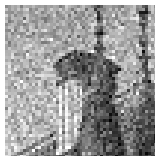
Inpainting

Заключение

Image denoising: сравнение результатов



original



noisy ($\sigma = 25$)



FoE



Portilla et al.



on-local means



non-linear diffusion

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Image denoising: результаты FoE



$\sigma = 15$, PSNR: 30.22 dB

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Image inpainting



Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

$$p(\mathbf{y}|\mathbf{x}) = \prod_{k=1}^{L \cdot M} p(y_k|x_k),$$

$$p(y_k|x_k) = \begin{cases} 1, & k \in \mathcal{M} \\ \delta(y_k - x_k), & k \notin \mathcal{M} \end{cases}$$

$$\mathbf{x}^{(t+1)} = \mathbf{x}^{(t)} + \eta \mathbf{M} \left[\sum_{i=1}^N \mathbf{J}_-^{(i)} * \psi'(\mathbf{J}^{(i)} * \mathbf{x}^{(t)}; \alpha_i) \right]$$

\mathcal{M} — мн-во пикселей, помеченых для заполнения

\mathbf{M} — соответствующая бинарная маска

- ▶ Такой метод не подходит для заполнения сложных текстур

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Image inpainting: сравнение методов I



Зарисованное текстом

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Image inpainting: сравнение методов I



Предложенный метод с применением FoE
(PSNR 29.06dB)

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Image inpainting: сравнение методов I



Bertalmío et al.
(PSNR 27.56dB)

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Image inpainting: сравнение методов II



Предложенный метод с применением FoE

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

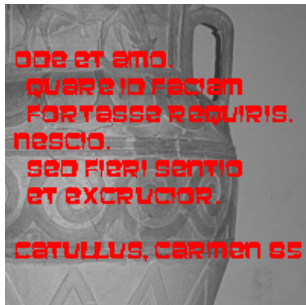
Приложения

Denoising

Inpainting

Заключение

Image inpainting: результаты FoE



Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

Image inpainting: результаты FoE

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

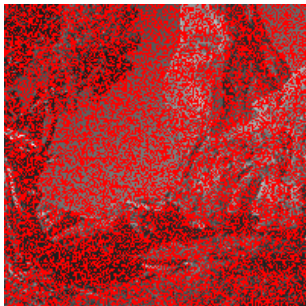
Обучение

Приложения

Denoising

Inpainting

Заключение



Outline

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Image denoising

Image inpainting

Заключение

Fields of Experts

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение

- ▶ Фильтры 5×5 достаточно малы, чтобы описывать текстурированные области. Небольшое увеличение фильтров (эксперимент 7×7) не особо помогает. Значительное увеличение затратно.
 - ▶ Сэмплы из $p_{\text{FoE}}(\mathbf{x})$ не выглядят натуральными.
 - ▶ Размытые области еще сильнее размываются, шум на текстурах убирается не полностью.
- ▶ Модель зависит от масштаба изображений (у конкурентов, кстати, не зависит).
- ▶ Размер клика можно было бы настраивать автоматически.
- ▶ Градиентный вывод применим не всегда, нужны исследования техник вывода для high-order MRF.

Введение

Image priors

Модели патчей

Fields of Experts

Модель

Вывод

Обучение

Приложения

Denoising

Inpainting

Заключение