Компьютерное зрение '2014

Who? Александр Вахитов

When? April 25, 2014

План лекции

Трекинг объекта

Линейная модель динамической системы

Сегментация изображений

Основные понятия

Байесовский подход

Поиск минимума

Примеры

Геометрия

3х мерная геометрия

Проективная камера

Изображение плоскости

Модель линейной эволюции

$$x_n = A_n x_{n-1} + w_n$$

x_n - вектор состояния системы, n - момент времени, $w_n \in N(0, \Sigma_w)$ - гауссовский вектор помехи стостояния

Пример

Движение точки с координатами (i, j) с постоянной скоростью (v_i, v_j):

$$x_n = (i, j, v_i, v_j)^T, \; A_n = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Модель наблюдения

$$z_n = H_n x_n + v_n,$$

z_n - вектор наблюдений (измерений), $v_n \in N(0, \Sigma_v)$ - гауссовский вектор помехи наблюдения, у нас $z_n = x_n + v_n$

Основные понятия

Наблюдаемость - можно ли по наблюдениям восстановить состояния?
Наблюдаемость эквивалентна $rank\ (H, HA, H^2xA, \ldots, H^{q-1}A)^T = q$, где q - размерность состояния.

Фильтр Калмана

Обозначим $P_{n|k} = cov(x_{n|k}, x_{n|k})$ Экстраполяция:

$$\hat{x}_{n|n-1} = A_n \hat{x}_{n-1|n-1}, \; P_{n|n-1} = A_n P_{n-1|n-1} A_n^T + \Sigma_w$$

Коррекция:

$$y_n = z_n - H_n \hat{x}_{n|n-1}, \; S_n = H_n P_{n|n-1} H_n^T + \Sigma_v,$$

$$K_n = P_{n|n-1} H_n^T S_n^{-1}$$

$$\hat{x}_{n|n} = \hat{x}_{n|n-1} + K_n y_n$$

$$P_{n|n} = (I - K_n H_n) P_{n|n-1}$$

Начальные данные: $\hat{x}_0$ - начальное приближение, трактуемое как случайная величина, $P_{0|0}$ - ковариация ошибки начального приближения

Основной источник

Simon Prince. Computer Vision Models, Learning, and Inference. 2012

Марковское случайное поле

Определение

Марковское случайное поле (MRF) - тройка:

- множество позиций (напр., пикселей)

$$S = \{1 \dots N\}$$

- множество случайных величин w_n, по одному для каждой позиции

$$\{w_n\}_{n=1}^{N}$$

- множества соседей $\{\mathcal{N}_n\}_{n=1}^{N}$ для каждой позиции

$$m \in \mathcal{N}_n \Leftrightarrow (m, n) \in C$$

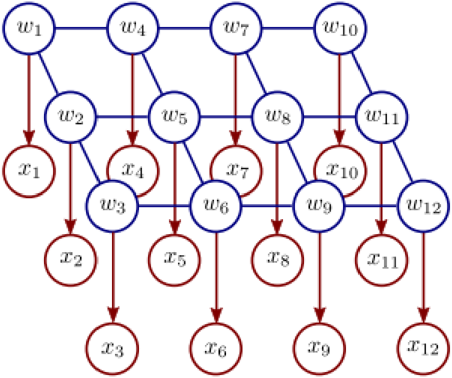
Основное свойство:

$$P(w_n | \mathbf{w}_{S \setminus n}) = P(w_n | w_{\mathcal{N}_n})$$

Наблюдения и параметры

Пусть в марковском случайном поле к каждой позиции привязан неизвестный случайный параметр w_n и наблюдение x_n .

Например, w_n - истинное значение пикселя, x_n - наблюдаемое значение, искаженное шумом.



Формула Байеса

Обозначим $\mathbf{w}$ - вектор параметров, $\mathbf{x}$ - вектор наблюдений.
Формула Байеса:

$$P(\mathbf{w}|\mathbf{x}) = \frac{P(\mathbf{x}|\mathbf{w})P(\mathbf{w})}{P(\mathbf{x})}$$

При этом,

$$argmax_{\mathbf{w}} P(\mathbf{w}|\mathbf{x}) = argmax_{\mathbf{w}} P(\mathbf{x}|\mathbf{w})P(\mathbf{w})$$

Байесовский подход к марковскому полю

Потенциальная функция $\phi_j(\cdot)$ - вероятность совместного появления значений на связанных позициях марковского поля. Сопоставим потенциальную функцию каждой позиции, при этом

$$\phi_j(w_n) = P(\mathbf{w}_{\mathcal{N}_n})$$

Тогда априорная вероятность $P(\mathbf{w})$ формулируется как

$$P(\mathbf{w}) = \Pi_{n=1}^{N} \phi_n(\mathbf{w}_{\mathcal{N}_\mathbf{n}})$$

В то же время,

$$P(\mathbf{x}|\mathbf{w}) = \Pi_{n=1}^{N} P(x_n|w_n)$$

Пример марковского поля

4 пикселя 1,2,3,4,

$$C = \{(1,2),\ (2,3),\ (3,4),\ (4,1)\}$$

$w_{1...4}$	$Pr(w_{1...4})$	$w_{1...4}$	$Pr(w_{1...4})$
0000	0.47176	0100	0.00471
0001	0.00471	0101	0.00005
0010	0.00471	0110	0.00471
0011	0.00471	0111	0.00471

$w_{1...4}$	$Pr(w_{1...4})$	$w_{1...4}$	$Pr(w_{1...4})$
1000	0.00471	1100	0.00471
1001	0.00471	1101	0.00471
1010	0.00005	1110	0.00471
1011	0.00471	1111	0.47176

Логарифмирование вероятности

Ищем

$$\hat{w} = argmax\ P(\mathbf{w}|\mathbf{x})$$

Обозначим

$$U_n(w_n) = -log\ P(x_n|w_n),\ \psi(w_m, w_n) = -log\ P(w_m, w_n).$$

$$argmax\ P(\mathbf{w}|\mathbf{x}) = argmax\ \Pi_{n=1}^{N}\{P(x_n|w_n)\}\Pi_{n=1}^{N}\phi_n(\mathbf{w}_{\mathcal{N}_\mathbf{n}})$$

$$= argmax \sum_{n=1}^{N} log\ P(x_n|w_n) - \sum_{(m,n)\in C} \psi(w_m, w_n) =$$

$$= argmin \sum_{n=1}^{N} U_n(w_n) + \sum_{(m,n)\in C} \psi(w_m, w_n).$$

Минимальный разрез и максимальный поток

Пусть имеется направленный ациклический взвешенный граф. Пусть в нем выделены вершина-исток и вершина-сток. Разрезом называется набор ребер, при удалении которых более не существует путь из истока в сток.

Трактуем вес ребра как пропускную способность.

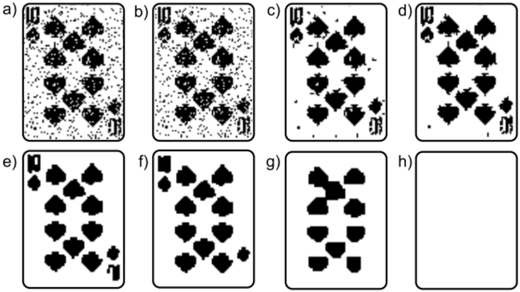
Для каждого пути из истока в сток пропускная способность есть минимум из пропускных способностей ребер в пути. В каждом пути есть ребро, на котором достигается минимум (насыщенное).

Разрезая последовательно насыщенные ребра, получим разрез графа. Разрез будет минимальным по сумме весов ребер, в него входящих.

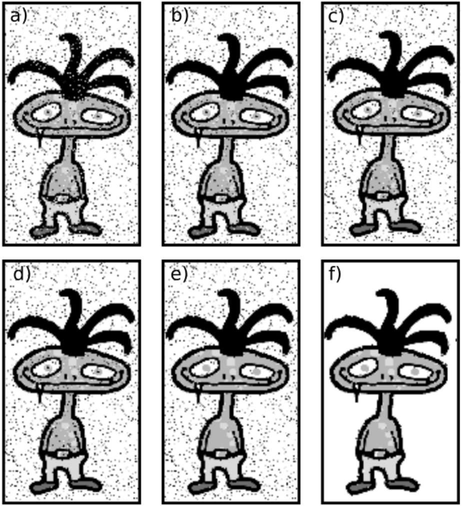
Поиск минимума

Для простоты, считаем w_n распределенным по Бернулли (принимает 0 или 1 с вероятностями ρ и $1-\rho$).
Пусть задан граф $\mathbf{G} = (\mathcal{V}, \mathcal{E})$. Зададим специальную вершину - исток и вершину - сток, а также N вершин (по числу позиций).
Соединим ребрами вершины m, n если $(m, n) \in C$.
Если отношение симметрично, получим ненаправленный граф.
Каждому ребру - вес: для ребер из истока $U_n(0)$, для ребер в сток $U_n(1)$, для ребра m, n - $\psi(m, n)$.
Найдем минимальный разрез - это и будет решение задачи.

Пример: удаление шумов



Пример: удаление шумов, метод расширения



Реальные приложения

- Стереозрение
- Сегментация (фон-передний план, или сложнее)
- Разность с фоном

3D геометрия

$$x \in \mathbb{R}^3$$

Евклидово преобразование - сохраняет расстояние между всеми точками.
Нетрудно получить, что любое евклидово преобразование - это

$$x' = Rx + t,$$

где $R \in M_{3\times 3}$ - матрица, $|detR| = 1$, ортогональная, $t \in \mathbb{R}^3$
При этом, если $det\, R = 1$, то это матрица поворота, иначе - поворота с отражением.

Параметризация вращений

Формула Родрига задает результат поворота вектора v на угол θ вокруг оси k, $\|k\| = 1$

$$v_{rot} = vcos\theta + (k \times v)sin\theta + k(k \cdot v)(1 - cos\theta)$$

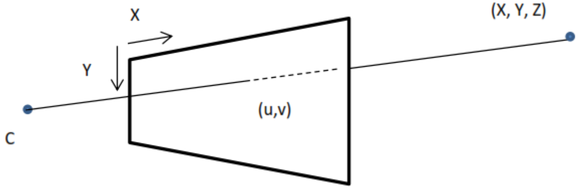
Позволяет получить по оси и углу матрицу поворота и обратно
Ось и угол кодируются в $\mathbb{R}^3$ как

$$\hat{\theta} = k\theta \in \mathbb{R}^3$$

Другие преобразования

Аффинное: наклон, растяжение
Проективное

Проективная камера с единичным фокусом



Начало координат - точка C, фокусное расстояние (от C до плоскости изображения) равно 1

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}, \ \lambda = \frac{1}{Z}$$

Проективная камера (2)

Пиксельные координаты в плоскости изображения сонаправлены с осями CX, CY. Начало отсчета пиксельных координат O_I, проекция C на плоскость изображения (c_x, c_y). Пиксель квадратный и имеет размер $p \times p$, тогда фокус в пиксельных единицах $f_p = p^{-1} f$. Пусть фокус равен f.

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = \begin{pmatrix} p^{-1} u + c_x \\ p^{-1} v + c_y \\ 1 \end{pmatrix} =$$

$$= \lambda \begin{pmatrix} f_p & 0 & c_x \\ 0 & f_p & c_y \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

Однородные координаты

$$\mathbf{x} \in \mathbb{R}^n \implies \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix} \in \mathbb{R}^{n+1}$$

Определение

Отношение эквивалентности:

$$\forall a, b \in \mathbb{R}^{n+1} : \; a \equiv b \Leftrightarrow \exists k \in \mathbb{R}^n : \; a = kb$$

Проективное пространство $\mathbb{P}^n$ определеяется как фактор-множество

$$\mathbb{P}^n = \mathbb{R}^{n+1} / \equiv$$

Пример

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \in \mathbb{P}^2$$

Евклидово преобразование в однородных координатах

$$\mathbf{x}' = R\mathbf{x} + \mathbf{t}$$

$$\begin{pmatrix} \mathbf{x}' \\ 1 \end{pmatrix} = (R \mid \mathbf{t}) \begin{pmatrix} \mathbf{x} \\ 1 \end{pmatrix}$$

Если координаты точки заданы в глобальной системе $OX_0Y_0Z_0$, и евклидово преобразование от $OX_0Y_0Z_0$ к $CXYZ$ есть $R, \mathbf{t}$, то

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \equiv \begin{pmatrix} f_p & 0 & c_x \\ 0 & f_p & c_y \\ 0 & 0 & 1 \end{pmatrix} (R \mid \mathbf{t}) \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Матрица камеры

$$K = \begin{pmatrix} f_p & 0 & c_x \\ 0 & f_p & c_y \\ 0 & 0 & 1 \end{pmatrix}$$

$$P = K\,(R \mid \mathbf{t})$$

$$\begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \equiv P \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Внутренняя калибровка камеры: f_p, c_x, c_y, внешняя калибровка камеры: $R, \mathbf{t}$

Поправка на дисторсию

Коэффициенты радиальной дисторсии
$K_1, K_2, K_3 \in \mathbb{R}$;
тангенциальной дисторсии $P_1, P_2, P_3 \in \mathbb{R}$

$$\bar{x} = x - c_x,\ \bar{y} = y - c_y,$$

$$r^2 = \bar{x}^2 + \bar{y}^2,$$

$$x' = x + \bar{x}(K_1 r^2 + K_2 r^4 + K_3 r^6 + \ldots) +$$

$$+[P_1(r^2 + 2\bar{x}^2 + 2P_2\bar{x}\bar{y}][1 + P_3 r^2 + \ldots],$$

$$y' = y + \bar{y}(K_1 r^2 + K_2 r^4 + K_3 r^6 + \ldots) +$$

$$+[2P_1\bar{x}\bar{y} + P_2(r^2 + 2\bar{y}^2)][1 + P_3 r^2 + \ldots],$$

Примеры: проекция с искажением

Дисторсированное (слева) и проективное (справа) изображения

В реальности, дисторсия зависит от расстояния до объекта и фокуса в метрах (для различных уровней зума - отличается дисторсия)

Точка на бесконечности

Точки прямой:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} X_a \\ Y_a \\ Z_a \end{pmatrix} + t \begin{pmatrix} X_b \\ Y_b \\ Z_b \end{pmatrix}$$

Изображение точек прямой

$$f\frac{X}{Z} = f\frac{X_a + tX_b}{Z_a + tZ_b} \to_{t \to \mathrm{inf}} f\frac{X_b}{Z_b}$$

Не зависит от $X_a, Z_a \implies$ параллельные прямые сходятся на изображении

Исключения?

Точка на бесконечности в проективных координатах

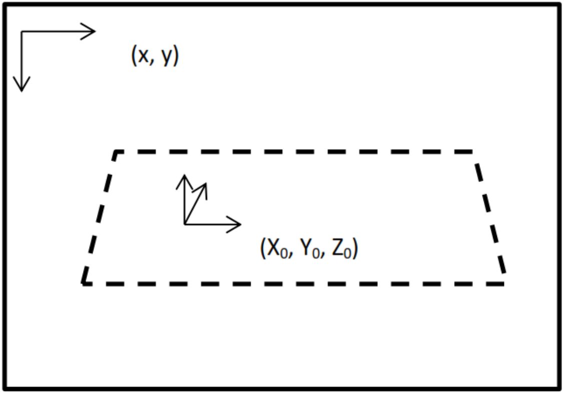
Точки прямой:

$$\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} = \begin{pmatrix} X_a + tX_b \\ Y_a + tY_b \\ Z_a + tZ_b \\ 1 \end{pmatrix} \equiv$$

$$\equiv \begin{pmatrix} X_a/t + X_b \\ Y_a/t + Y_b \\ Z_a/t + Z_b \\ 1/t \end{pmatrix} \rightarrow_{t\rightarrow \mathrm{inf}} \begin{pmatrix} X_b \\ Y_b \\ Z_b \\ 0 \end{pmatrix}$$

Например, прямая с $X_b = 1,\ Y_b = 0, Z_b = 0$ сходится к $(1, 0, 0, 0)^T$

Изображение плоскости



$$R = (\mathbf{r_1} \mid \mathbf{r_2} \mid \mathbf{r_3}); \quad r_1 = (R \mid \mathbf{t}) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Изображение плоскости (2)

$$r_1 = (R \mid \mathbf{t}) \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad v_1 \equiv K r_1 = P \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Как по матрице P определить v_1?

Определение

v_1, v_2, v_3 - исчезающие точки, точки на бесконечности (зенита или надира для ортогонального плоскости направления, горизонта - для направления в плоскости)