ПОРІВНЯЛЬНИЙ АНАЛІЗ АЛГОРИТМІВ СТЕРЕОЗОРУ В РАМКАХ БАЙЄСІВСЬКОЇ ТЕОРІЇ ПРИЙНЯТТЯ РІШЕНЬ

Савчинський Б.Д.
Міжнародний науково-навчальний центр ЮНЕСКО/МПІ інформаційних технологій та систем.
пр.Академіка Глушкова 40, Київ 03680 т.266-62-08 bogdan@image.kiev.ua

Анотація. Розглядається байєсівський підхід до задачі стереозору. Існуючі її розв'язки аналізуються з точки зору байєсівських вирішуючих правил. Проводиться порівняльний аналіз алгоритмів стереозору в рамках байєсівського підходу.

1 ВСТУП

Проблема машинного стереозору має свою історію, що триває вже кілька десятиріч. Дослідження цієї проблеми були започатковані видатними результатами, що зводили задачі стереозору до вирішення певних задач динамічного програмування (див.,напр. [5], [6]). В рамках цього підходу стало можливим знаходити найімовірніший рельєф місцевості, що відповідає заданій стереопарі знімків. В цьому напрямі були вирішені численні прикладні проблеми. Завдяки одержаним успіхам такий підхід став свого роду аксіомою, і задачі, що не вдавалося вирішити у рамках цього підходу, вважалися за практично невирішувальні.

Дана робота відходить від встановленого стереотипу знаходження **найімовірнішого** рельєфу, що конче є лиш окремим випадком байєсівської теорії прийняття рішень. Погляд на задачу з точки зору цієї, більш загальної теорії, дозволив знайти нові, прийнятніші формулювання задачі стереозору і, відповідно, побудувати нові алгоритми її вирішення.

Робота складається з чотирьох розділів. Другий розділ присвячений постановці і розв'язку задачі стереозору, як задачі, що формулюється в рамках байєсівського підходу. Тут запропоновано різні функції втрат та алгоритми розв'язку задач, що їм відповідають. Третій розділ має на меті демонстрацію прикладних результатів, які були отримані при застосуванні алгоритмів, приведених у другому розділі. Нарешті, в четвертому розділі формулюються висновки.

2 ПОСТАНОВКА ТА РОЗВ'ЯЗОК ЗАДАЧІ СТЕРЕОЗОРУ ПРИ РІЗНИХ ФУНКЦІЯХ ВТРАТ

Задача стереозору полягає у відновленні просторової конфігурації об'єкта за парою його фотознімків. Вважатимемо, що об'єкт знімається таким чином, що для кожної його точки справедливо: або її зображення присутнє на обох знімках, або на жодному із знімків його нема. Зображення точки об'єкта на лівому та правому знімках називатимемо відповідно лівою та правою проекціями цієї точки. Не втрачаючи загальності вважатимемо, що зображення оцифровані і відкаліброванi таким чином, що рядки з номерами k на лівому та правому зображеннях є, відповідно, лівою та правою проекціями деякої лінії на вихідному об'єкті для будь-якого k. Отже, інформацію про просторову конфігурацію цієї лінії несуть лише дані два рядки. Тобто, ми можемо обмежитись розглядом задачі стереозору лише для деякої пари рядків.

Назвемо **полем зору** множину:

$$T = \{i : 0 \le i \le I\}.$$

Елементами поля зору є пікселi, або **точки**. **Зображенням** називатимемо функцію:

$$x : T \to R.$$

Множину всіх зображень позначатимемо через R^T. Значення зображення $x(t)$ в точці t називахться **кольором піксела**. Нехай x і y – два зображення. Введемо функцію $\delta : R \times R \to R$, яку називатимемо **різницею кольорів пікселів**. Окремими її випадками є $\delta(x(t_1), y(t_2)) =\mid x(t_1) - y(t_2) \mid$ та $\delta(x(t_1), y(t_2)) =\mid x(t_1) - y(t_2) \mid^2$.

Також розглянемо клас Φ монотонних функцій вигляду $\phi : T \to T$, тобто:

$$\Phi = \{\phi : t_1 > t_2 \Rightarrow \phi(t_1) \ge \phi(t_2),\ t_1, t_2 \in T\} \qquad (1)$$

Дана властивість означах, що *"точкам, які знаходяться правіше на лівому зображенні, повинні відповідати лише точки, які знаходяться правіше і на правому зображенні"*

.

При даному фіксованому ϕ точки t та $\phi(t)$ називатимемо **відповідними**. Для зручності введемо клас $\mathcal{P}$ функцій вигляду $\rho : T \to T$, який визначимо як:

$$\mathcal{P} = \{\rho : \exists \phi \in \Phi \ \forall t \in T \ \rho(t) = t - \phi(t)\}.$$

Функції з класу $\mathcal{P}$ називатимемо **функціями паралакса**, а значення $\rho(t),\ \rho \in \mathcal{P}$ - **паралаксами**. Очевидним є те, що задачі знаходження $\phi \in \Phi$ та $\rho \in \mathcal{P}$ є еквівалентними. Але паралакси мають прозорий геометричний зміст: висота об'єкта є монотонною фунцією паралакса.

’ўҐа¤Ґп 2.1 *Нехай* $\rho \in \mathcal{P}$*. Тоді множина значень, які може приймати функція* ρ *в точці* $t,\ 1 \le t \le I$ *при фіксованих значеннях* $\rho(0), \rho(1), ..., \rho(t-1)$ *залежить лише від* $\rho(t-1)$ *і не залежить від* $\rho(0), \rho(1), ..., \rho(t-2)$.

$\triangleright$ Доведення випливає із справедливості даного твердження для функцій з класу Φ, що, в свою чергу, є наслідком того, що $\phi \in \Phi \Leftrightarrow \phi(t-1) \le \phi(t)\ \forall t : 1 \le t \le I$. $\triangleleft$

Позначимо через $N_i(t),\ 0 \le i \le (I-1)$ множину тих значень, які може приймати функція ρ в точці $i+1$ при фіксованому значенні $\rho(i) = t$:

$$N_i(t) = \{r : \exists \rho \in \mathcal{P}\ \rho(i) = t \ \&\ \rho(i+1) = r\}.$$

Для формулювання задачі в рамках байєсівської теорії (див.,напр. [1],[2],[3]) нам необхідно визначити **множину ознак об'єкта**, **множину станів**, їх сумісний розподіл ймовірностей та **функцію втрат**.

Множина ознак об'єкта

$$X = \{(x_l, x_r) : x_l, x_r \in R^T\}$$

- це пара зображень. Множина станів збігається з множиною функцій паралакса $\mathcal{P} \ni \rho$. Вважатимемо, що апріорний розподіл $p(\rho)$ всіх можливих функцій паралакса є рівномірним, а умовний розподіл ознак об'єкта при заданому його стані визначається як:

$$p((x_l, x_r)/\rho) = A \cdot \prod_{t \in T} \exp[-\delta(x_l(t), x_r(\phi(t)))] =$$
$$= A \cdot \prod_{i=0}^{I} \exp[-\delta(x_l(i), x_r(i - \rho(i)))] =$$
$$= A \cdot \exp[-\sum_{i=0}^{I} \delta(x_l(i), x_r(i - \rho(i)))],$$

де A – нормувальний множник.

За таких припущень апостеріорний розподіл функцій паралакса при заданій парі зображень матиме вигляд:

$$p(\rho/(x_l, x_r)) = p((x_l, x_r), \rho)/p((x_l, x_r)) =$$
$$= p((x_l, x_r)/\rho)p(\rho)/p((x_l, x_r)) = B \cdot p((x_l, x_r)/\rho) =$$
$$= C \cdot \exp[-\sum_{i=0}^{I} \delta(x_l(i), x_r(i - \rho(i)))],$$

де C та B – нормувальні множники.

При заданій функції втрат $W : \mathcal{P} \times \mathcal{P} \to R$ задача полягах в знаходженні такої функції паралакса ρ^*, яка б мінімізувала умовне математичне сподівання втрат:

$$\rho^* = \arg\min_{q \in \mathcal{P}} \sum_{\rho \in \mathcal{P}} p(\rho/(x_l, x_r)) W(q, \rho). \qquad (2)$$

Розглянемо кілька різних функцій втрат:

$$W(q, \rho) = \begin{cases} 0, & \text{якщо } \rho = q \\ 1, & \text{якщо } \rho \ne q \end{cases} \qquad (3)$$

$$W(q, \rho) = \sum_{\substack{t \in T \\ \rho(i) \ne q(i)}} 1 \qquad (4)$$

$$W(q, \rho) = \sum_{t \in T} (\rho(t) - q(t))^2 \qquad (5)$$

Почнемо з (3). Згідно з (2)

$$\rho^* = \arg\min_{q \in \mathcal{P}} \sum_{\substack{\rho \in \mathcal{P} \\ \rho \ne q}} p(\rho/(x_l, x_r)) = \qquad (6)$$
$$= \arg\min_{q \in \mathcal{P}} (1 - p(q/(x_l, x_r))) =$$
$$= \arg\max_{q \in \mathcal{P}} p(q/(x_l, x_r)) = \qquad (7)$$
$$= \arg\max_{q \in \mathcal{P}} \exp[-\sum_{i=0}^{I} \delta(x_l(i), x_r(i - q(i)))] =$$
$$= \arg\min_{q \in \mathcal{P}} \sum_{i=0}^{I} \delta(x_l(i), x_r(i - q(i))). \qquad (8)$$

Отже, функції (3) відповідає оцінка максимальної правдоподібності (див. (7)) функції ρ. З іншого боку, **формула (8) виражає постановку задачі стереозору при традиційному, небайєсівському підході**. Але чи доречною є в поставленій задачі функція втрат вигляду (3) ? При такій функції залишаються неврахованими кілька факторів. По-перше, дана функція втрат не враховує **в скількох точках** відрізняються різні функції паралакса між собою. Неформально кажучи, ті функції паралакса, які відрізняються лише в одній точці, мусять штрафуватись менше, аніж ті, що відрізняються в десятьох точках. Цей факт враховується функціхю втрат (4). Та, більше того, оскільки паралакси монотонно пов'язані з висотою, то грає роль ще й **величина різниці** між функціями паралакса в тих точках, в яких вони відрізняються. Цей факт може бути врахований за допомогою функції втрат (5).

Але повернімось до розв'язку задачі (8). Введемо позначення:

$$F_k(\rho(0),\rho(1),...,\rho(k)) = \sum_{i=0}^{k} \delta(x_l(i), x_r(i-\rho(i))).$$

Між цими величинами існух рекурентне співвідношення:

$$\begin{aligned} F_k(\rho(0),\rho(1),...,\rho(k)) = \\ = F_{k-1}(\rho(0),\rho(1),...,\rho(k-1)) + \\ +\delta(x_l(k), x_r(i-\rho(k))), 1 \le k \le I, \end{aligned} \quad (9)$$

а, отже:

$$\min_{\substack{\rho(0),\rho(1),...,\rho(k)\\ \rho\in\mathcal{P}}} F_k(\rho(0),\rho(1),...,\rho(k)) =$$
$$= \min_{\substack{\rho(0),\rho(1),...,\rho(k-1)\\ \rho\in\mathcal{P}}} F_{k-1}(\rho(0),\rho(1),...,\rho(k-1)) +$$
$$+ \min_{\rho(k)\in N_{k-1}(\rho(k-1))} \delta(x_l(k), x_r(i-\rho(k))), 1 \le k \le I, \quad (10)$$
$$\min_{\rho(0)\in T} F_0(\rho(0)) = \min_{\rho(0)\in T} \delta(x_l(0), x_r(0-\rho(0))). \quad (11)$$

Оскільки:

$$\rho^* = \arg\min_{\rho\in\mathcal{P}} F_I(\rho(0),...,\rho(I)).$$

– то алгоритм розв'язку задачі (8) полягає в послідовному застосуванні формул (10) та (11) для $k = I; I-1; ...; 0$.

Перейдемо до функцій втрат, які задаються виразами (4) та (5). Їх можна представити у вигляді:

$$W(q,\rho) = \sum_{i=0}^{I} \omega(q(i),\rho(i)),\ q,\rho\in\mathcal{P}. \quad (12)$$

Функції втрат вигляду (4) відповідах

$$\omega(k,r) = \begin{cases} 0, & \text{якщо } r = k \\ 1, & \text{якщо } r \ne k. \end{cases}$$

А функції втрат вигляду (5) –

$$\omega(k,r) = (k-r)^2.$$

Тому розв'яжемо задачу (2) саме для такого, більш загального вигляду функцій втрат (12) :

$$\rho^* = \arg\min_{q\in\mathcal{P}} \sum_{\rho\in\mathcal{P}} p(\rho/(x_l,x_r)) \sum_{i=0}^{I} \omega(q(i),\rho(i)) =$$
$$= \arg\min_{q\in\mathcal{P}} \sum_{i=0}^{I} \sum_{\rho(i)\in T} p(\rho(i)/(x_l,x_r))\omega(q(i),\rho(i)), \quad (13)$$
$$\text{де } p(\rho(i)/(x_l,x_r)) = \sum_{\rho\in\mathcal{P}} p(\rho/(x_l,x_r),\rho(i))$$

– сумарна ймовірність всіх функцій паралакса ρ при заданому значенні $\rho(i)$ в точці i. Припустимо, що величини $p(\rho(i)/(x_l,x_r))$ нам відомі. Тоді задача (13) перетворюється в задачу (8) заміною величин $\delta(x_l(i), x_r(i-q(i)))$ на $\sum_{\rho(i)\in T} p(\rho(i)/(x_l,x_r))\omega(q(i),\rho(i))$, а, отже, і розв'язухться за допомогою співвідношень, що аналогічні співвідношенням (10) та (11).

Знаходження величин $p(\rho(i)/(x_l,x_r))$ також проводиться за тією ж схемою (10),(11). Але замість операції max необхідно підставити операцію $\sum$, а замість $\sum$ – операцію $\prod$. Запишемо необхідні перетворення:

$$p(\rho(i)=t/(x_l,x_r)) = \sum_{\rho\in\mathcal{P}} p(\rho/(x_l,x_r),\rho(i)=t) =$$
$$= C \sum_{\rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}} \prod_{j=0}^{I} exp[-\delta(x_l(j),x_r(j-\rho(j)))] =$$
$$= C\cdot \sum_{\rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}} \prod_{j=0}^{i} exp[-\delta(x_l(j),x_r(j-\rho(j)))] \times$$
$$\times \prod_{j=i}^{I} exp[-\delta(x_l(j),x_r(j-\rho(j)))] =$$
$$= C\left(\sum_{\substack{(\rho(0),\rho(1),...,\rho(i-1))\\ \rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}}} \prod_{j=0}^{i} \exp[-\delta(x_l(j),x_r(j-\rho(j)))]\right) \times$$
$$\times\left(\sum_{\substack{(\rho(i+1),...,\rho(I))\\ \rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}}} \prod_{j=i}^{I} \exp[-\delta(x_l(j),x_r(j-\rho(j)))]\right) \quad (14)$$

Обчислимо перший співмножник у виразі (14). Другий обчислюється аналогічним чином. Процедура обчислення має вигляд:

$$F_k(\rho(0),\rho(1),...,\rho(k)) \stackrel{def}{=}$$
$$\stackrel{def}{=} \prod_{j=0}^{k} \exp[-\delta(x_l(j),x_r(j-\rho(j)))]; \quad (15)$$

$$\begin{aligned} \Rightarrow F_k(\rho(0),\rho(1),...,\rho(k)) = \\ = F_{k-1}(\rho(0),\rho(1),...,\rho(k-1)) \times \\ \times \exp[-\delta(x_l(k),x_r(j-\rho(k)))]; \end{aligned} \quad (16)$$

$$\Rightarrow \sum_{\substack{(\rho(0),\rho(1),...,\rho(k))\\ \rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}}} F_k(\rho(0),\rho(1),...,\rho(k)) =$$
$$= \sum_{\substack{(\rho(0),\rho(1),...,\rho(k-1))\\ \rho\in\mathcal{P}\cap\{\rho:\rho(i)=t\}}} F_{k-1}(\rho(0),\rho(1),...,\rho(k-1)) \times$$
$$\times \sum_{\rho(k)\in N_{k-1}(\rho(k-1))} \exp[-\delta(x_l(k),x_r(k-\rho(k)))] \quad (17)$$

$$F_0(\rho(0)) = \sum_{\rho(0)\in T} \exp[-\delta(x_l(0), x_r(-\rho(0)))] \quad (18)$$

Таким чином, алгоритм розв'язку пошуку першого співмножника у виразі (14) полягає в послідовному застосуванні формул (17) та (18) для $k = i; i-1; ...; 0$.

3 ДЕМОНСТРАЦІЯ ПРИКЛАДНИХ РЕЗУЛЬТАТІВ

В якості вихідних даних (стереопари) було взято відкалібровану стереопару обличчя людини (див. [4]).

На рис.1 зображено результат відновлення рельєфу при трьох різних функціях втрат.

Результати представлені у вигляді так званого "висотного"зображення, тобто більша яскравість точки відповідах її більшій висоті. Справа і знизу від зображення показані "розрізи": вертикальний та горизонтальний. Найлівіше зображення відповідає функції вигляду (3), зображення по центру – (4) і найправіше – (5).

Різниця в результатах є очевидною. Слід все ж відзначити, що при традиційному, небайєсівському підході (при якому використовуються критерії ,аналогічні (8)), можна отримати результати, що можуть бути порівняні з результатами, отриманими в рамках байєсівського підходу з функціхю втрат (5) . Це досягається головним чином за рахунок ускладнення моделі (тобто структури класу Ф) та використання різних евристичних прийомів.

4 ВИСНОВКИ

- В даній роботі продемонстровано відновлення рельєфу на основі байєсівського підходу при трьох різних функціях втрат. При цьому найкращі результати було отримано при інтуїтивно найкращій з них – функції втрат (5).
- Метод найбільшощ правдоподібності не повинен бути догмою. Байєсівський підхід включає його як частковий випадок і тому є більш потужним інструментом при прийнятті рішень.

Література

[1] Schlesinger M.I., V.Hlavač. Deset prednasek z teorie statistickeho a strukturniho rozpoznavani. Vydavatelstvi CVUT, Praha, 1999.

[2] Ковалевский В.А. Методы оптимальных решений в распознавании изображений.–М. :Наука, 1976.–328с.

[3] Дуда Р., Харт П. Распознавание образов и анализ сцен. – Москва: Мир, 1976.

[4] http:\\www.lincoln3d.com

[5] Гимельфарб Г.Л. Симметричный подход к задаче автоматических стереоскопических измерений в фотограммметрии. *Кибернетика*.- 1979,(2).-с.73-82.

[6] Гимельфарб Г.Л. Симметризованное би- и тринокулярное стереозрение: взаимосвязи между теоретическими основами и эвристическими решениями. *Теоретические и прикладные вопросы распознавания изображений*.- Киев, 1995.-с.4-25.

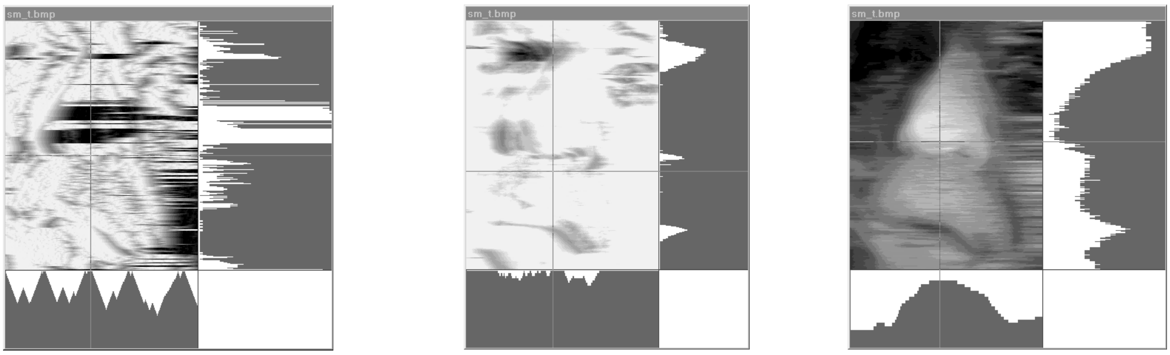

Рис. 1 Результати розв'язку задачі стереозору для трьох описаних випадків. Малюнок зліва відповідає пошуку найімовірнішого рельєфу (див.(8)), по центру - розв’язку байєсівської задачі з “точковим” штрафом (4). І, нарешті, малюнок справа отриманий з розв’язку байєсівської задачі з “квадратичним” штрафом (5).