

ВИКОРИСТАННЯ НЕЙРОМЕРЕЖЕВИХ СТРУКТУР ДЛЯ ПОБУДОВИ СИСТЕМ РОЗПІЗНАВАННЯ ОБРАЗІВ З КОМПЕНСАЦІЄЮ ІНФОРМАЦІЙНИХ ПОТОКІВ

Рожков С.О., Бражник Д.О.

Кафедра технічної кібернетики
Херсонський державний технічний університет
app@kstu.edu.ua

Анотація

Розглянуто задачу розпізнавання оптичних образів і сцен в умовах невизначеності умов пред'явлення і збурювань у просторі оригіналу. В основу роботи покладений метод компенсації вхідних інформаційних потоків [1].

Основна проблема методу полягає в реалізації вхідного перетворювача інформації, що дозволяє судити про відповідність вхідного сигналу сигналу еталону. Показано, що як датчик у системі можливо використання нейронної мережі з вхідним сигналом, що модулює матрицю ваг, при цьому вхідний вектор генерується як рядок еталона.

Вступ

Особливості сучасних систем розпізнавання образів полягають у тім, що вони будуються як системи перетворення інформації. Це визначає їхню чутливість до умов пред'явлення вхідних образів. З іншої сторони відомо, що інваріантність до збурювань забезпечують системи, побудовані з використанням зворотних зв'язків – системи по відхиленню [1]. У роботі розглянутий принцип компенсації вхідних інформаційних потоків, тобто, розглянута система, що компенсує вхідний вплив за рахунок генерації еталонних сигналів.

1. Використання схеми компенсації вхідних потоків інформації

Якщо розглядати потоки інформації образу J_0 і еталона J_e , як зустрічні з мірою відхилення J_e , де $J_e = J_e(\mathbf{x}, \mathbf{u}, \Omega_i, t)$ - умовна інформація J_{oe} , то можливо постановку задачі оптимального керування для еталона Ω_i : $\mathbf{x}^*, \mathbf{u}^* \rightarrow \min J_e$ рішення якої доставляє рішення задачі розпізнавання в класі образів Ω : $\Omega_i \rightarrow \inf J_e$.

На рис. 1 показана структура системи, отримана при такому підході. Як відхилення вхідного сигналу об'єкта від сигналу моделі виступає J_e -

умовна інформація зв'язку між потоком J_0 і потоком J_e .

Використовуючи ідею методу припустимих перетворень [2], визначимо зручний для спостереження параметр, однозначно зв'язаний з J_e і дотримуючись рекомендацій [3], розглянемо кореляційну міру близькості між еталонним ω_i^* і пред'явленим образом ω . При цьому апертурне обмеження [4], зв'язане з перевищенням геометричних розмірів полюси контролю над "вікном" датчика, визначає введення активного руху датчика, що сканує, в полі контролю.

Сигнали еталона $f^*(\mathbf{x})$ і об'єкта $f(\mathbf{x})$ залежать від закону руху сканування в просторі еталона як $\mathbf{x}_e = \mathbf{x}_e(t)$ і закону сканування в просторі об'єкта $\mathbf{x}_o = \mathbf{x}_o(t)$. Очевидно, при усередненні по реалізаціях, відстань між еталоном і об'єктом:

$$\begin{aligned} \rho^* &= \max_{\Delta \mathbf{x}} \rho(\omega_i, \omega_i^*, \Delta \mathbf{x}) = \\ &= \max_{\Delta \mathbf{x}} M \{ f^0(\mathbf{x}) * f^0(\mathbf{x}) \} = \\ &= \max_{\Delta \mathbf{x}} \iint_{S_\omega} f^0(\mathbf{x}) f^0(\mathbf{x} + \Delta \mathbf{x}) dS, \end{aligned}$$

де $f^0(\mathbf{x})$, $f^*(\mathbf{x})$ - центровані сигнали еталона й об'єкта, S_ω - область чутливості датчика, $\Delta \mathbf{x}$ - просторове зрушення між еталоном і зображенням.

Після усереднення за часом одержимо середнє значення відхилення об'єкта й еталона:

$$\begin{aligned} \overline{\rho^*} &= \max_{\Delta \mathbf{x}} \frac{1}{T} \int_0^T \rho^* dt = \\ &= \iint_{S_\omega} \left(\max_{\Delta \mathbf{x}} \frac{1}{T} \int_0^T f(\mathbf{x}_o(t)) \cdot f^*(\mathbf{x}_e(t) + \Delta \mathbf{x}) dt \right) dS \quad (1) \end{aligned}$$

Просторове зрушення $\Delta \mathbf{x}$ в області S_ω для сигналу об'єкта визначає рівнобіжне представлення зображення об'єкта у вікні S_ω . При цьому, за умови збігу об'єкта й еталона $f=f^*$ існує $\Delta \mathbf{x}^*$, при якому:

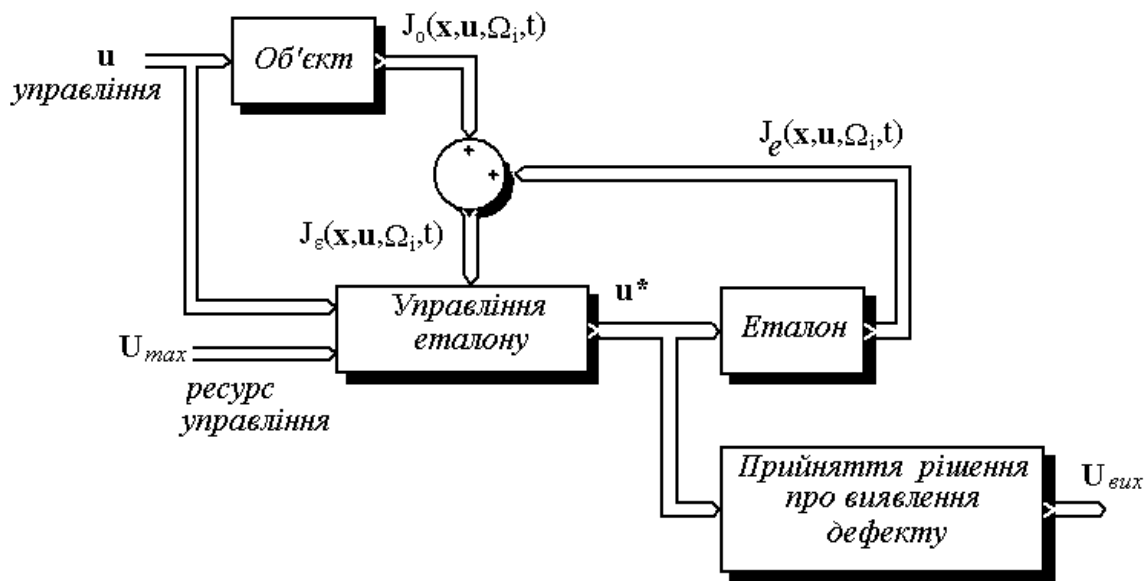


Рис.1 - Структурна схема системи оптимальної фільтрації

$$\Delta \mathbf{x}^* \rightarrow \sup_{\Delta \mathbf{x} \in S\omega} \left(\frac{1}{T} \int_0^T f(\mathbf{x}_o(t) + \Delta \mathbf{x}) \cdot f^*(\mathbf{x}_o(t)) dt \right) d. \quad (2)$$

$\Delta u = 0.$

У даній задачі внесена невизначеність сполучення в межах $S\omega$. Тому умова (2) переходить в умову $\Delta \mathbf{x} \in S\omega, \Delta \mathbf{x} = \text{const}$. Це дозволяє поставити двоїсту задачу при обмеженнях на керування еталонном $\Delta \mathbf{u}^*(t) \rightarrow \Delta \mathbf{x} = \text{const}$. Обмеження на керування $\Delta \mathbf{u}^*(t) \in \Omega$ визначається вимогою зв'язності околиці образу, тобто, якщо деформації не перевищують припустимих значень, то витрати на керування не повинні перевищувати припустимого ресурсу керування.

Таким чином, якщо задача оптимального керування (2) при обмеженнях $\Delta \mathbf{x} = \text{const}$ і $\Delta \mathbf{u}^*(t) \in \Omega$ має рішення, то образи, що породжують сигнали f і f^* , будуть близькими. При цьому, чим менше витрати на керування, тим ближче образи ω_i^* і ω_i .

Це дає алгоритм рішення задачі розпізнавання:

$$\omega_i^* \rightarrow \inf_{\mathbf{x} \in S\omega} \Delta \mathbf{u}^*. \quad (3)$$

Отже, виконання умов загасання керування, сталості координат кореляційного максимуму $\Delta \mathbf{x} = \text{const}$ і умова збігу еталона й об'єкта $f = f^*$ доставляють supp^* .

2. Використання нейронної мережі як чутливого елемента системи

Однак найбільшу складність при рішенні задачі викликає реалізація алгоритму визначення положення кореляційного максимуму. Для рішення даної задачі необхідні рівнобійні обчислення великого числа кореляційних функцій [5].

Найбільший ступінь паралельності для даного процесу забезпечує нейронна мережа, що має досить просту структуру (рис. 2).

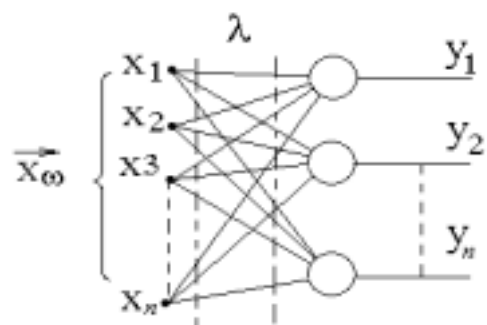


Рис. 2 - Проста нейронна мережа

Для простої мережі вектор вхідного сигналу перетворюється у вектор виходу в загальному випадку множенням на матрицю ваг, що зберігає інформацію про розпізнавані образи (4).

$$A\mathbf{x}_{\omega} = \mathbf{y} \quad \leftrightarrow \quad \omega_i \quad (4)$$

Природно, що розмірність вектора входу значно менше розмірності пам'яті $\dim \mathbf{x} = n$; $\dim A = n \times n$, де власне рядка матриці ваг являють собою набір еталонів розпізнаваних образів. Таким чином, структурно робота мережі представляє перетворення сигналу системою (рис. 3).

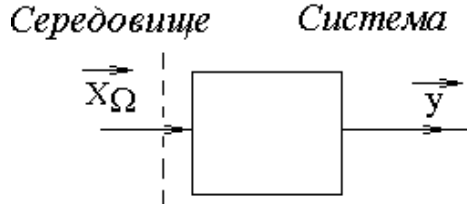


Рис. 3 - Взаємодія системи з зовнішнім середовищем

Даний підхід припускає високий ступінь паралельності обчислень, що реально можливо тільки при аналоговій обробці сигналу. Однак процес обчислення реалізується як послідовний і достоїнства алгоритму зникають.

З іншого боку, можна реалізувати матрицю фотоприймачів, що обробляє сигнал по такому ж алгоритму, з тією лише різницею, що вхідний образ відбиває матрицею ваг (5)

$$A_{\omega} \mathbf{x} = \mathbf{y}. \quad (5)$$

У цьому випадку вхідний вектор повинний генеруватися як рядок очікуваного зображення, а реакція визначається по компонентах вихідного вектора (6).

$$\sum_{i=1}^n a_{ij} x_i = y_j \quad (6)$$

Структура мережі в цьому випадку не змінюється, але вхідним сигналом є значення ваг (рис. 4).

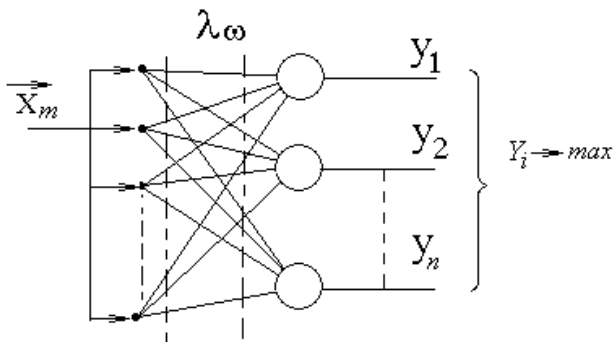


Рис. 4 - Мережа при вході по матриці ваг

Дуже істотно те, що мережа виступає як елемент порівняння в замкнутій системі, де мається сигнал із

системи, а вихідний вектор несе інформацію про ступінь збігу вхідного образу й еталона. Отже, взаємодія системи з зовнішнім середовищем представляє компенсацію зовнішнього потоку інформації і внутрішнього потоку інформації від еталона (рис.5).

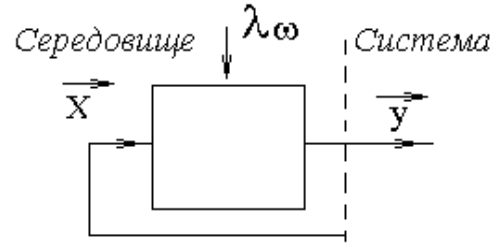


Рис.5 - Взаємодія системи та середовища з компенсацією інформації

Реалізація матриці вхідного образа, природно, допускає центрування на ковзне середнє. Так само можливе центрування і рядка еталону.

$$\bar{a}_{ij} = a_{ij} - \frac{1}{n} \sum_{i=1}^n \sum_{j=1}^n a_{ij}; \quad \bar{x}_i = x_i - \frac{1}{n} \sum_{i=1}^n x_i.$$

У такому випадку легко представити кожний з компонентів вихідного вектора як оцінку кореляції рядка вхідного образа й еталону (7).

$$\sum_{i=1}^n \bar{a}_{ij} \bar{x}_i = y_j \rightarrow y_i \approx R_{\omega x}. \quad (7)$$

Якщо у вхідному сигналі утримується рядок, що відповідає очікуваному образіві вхідного сигналу, то одержимо максимальне значення оцінки кореляції рядка, що підтверджує гіпотезу про відповідність вхідного образа еталона

$$\mathbf{x} \in \omega \rightarrow R_{\omega x} \rightarrow \max.$$

У даному випадку, природно, необхідне введення підсистеми керування, що забезпечує стабілізацію положення обраних рядків, як у горизонтальному напрямку, так і по вертикальному.

$$\begin{aligned} u_x, \quad \mathbf{x}(u_x), \quad u_x \rightarrow \max y_i, \\ i \neq \text{var}; \\ u_y, \quad \mathbf{x}(u_y), \quad u_y \rightarrow \max y_i. \end{aligned}$$

Введення обмежень на ресурс керувань забезпечує збереження зв'язності крапок вхідного образу й еталона.

$$(u_{x \max}, u_{y \max}), \quad y_i = \max,$$

$$\begin{cases} |u_x| < u_{x \max} \\ |u_y| < u_{y \max} \end{cases} \rightarrow \mathbf{x} \in \mathcal{W}.$$

Таким чином, при наявності підсистем стабілізації кореляційного максимуму можливо просто генерувати рядок образу, що являється гіпотезою і якщо ресурс керування достатній, можна затверджувати про несуперечності вхідного образу і висунутої гіпотези.

У даному випадку процес розпізнавання розбивається на три – висування гіпотези, природно з контексту, генерація еталонного сигналу і стабілізація кореляційного максимуму на фотоприймачі.

Істотним моментом у даному випадку є той факт, що дані процеси виконуються паралельно. Дійсно, після висування гіпотези, генерація еталона йде самостійно до повного завершення або до моменту втрати керування по перевищенню припустимого для даного образу ресурсу управління. Робота підсистеми стабілізації кореляційного максимуму не вимагає аналізу до моменту завершення генерації еталона або до моменту втрати керування. Робота підсистеми висування гіпотези не задіяна до кінця процедури перевірки. Але момент утрати керування зв'язаний з генерацією еталона і для коректування гіпотези надаються дані про області невідповідності еталона й об'єкта.

У випадку, коли матриця фотоприймачів не покриває весь об'єкт або вхідний образ є тільки частиною більш складного зображення $\mathcal{W} \subset \Omega$ природно ввести додатковий рух, що сканує у просторі вхідного образу, так і в просторі еталону

$$\mathcal{W}^* \Rightarrow \mathbf{x} = \mathbf{x}(\varphi_x(t), \varphi_y(t)) = \mathbf{x}(\varphi(t)).$$

При справедливості висунутої гіпотези керування скануванням вхідного простору і скануванням еталона розрізняються на керування необхідне для усунення перекручувань вхідного образу

$$\mathbf{u}_\Delta(t) = \mathbf{f} - \varphi.$$

Як і в попередньому випадку необхідне обмеження ресурсу керування, для забезпечення перевірки умови зв'язності крапок вхідного образу й еталона. Якщо на всій траєкторії виконується умова $\mathbf{u}_{\Delta \max} > \mathbf{u}_\Delta$, то між еталоном і образом немає порушень зв'язності й умови підтвердження гіпотези приймають вид

$$y_i = \max, \quad |\mathbf{u}_\Delta| < \mathbf{u}_{\Delta \max},$$

$$\begin{cases} |u_x| < u_{x \max} \\ |u_y| < u_{y \max} \end{cases} \rightarrow \mathcal{W}^* \in \Omega.$$

Для випадку складного зображення необхідна рівнобіжна робота декількох аналізаторів. У такому випадку і генерація еталонів ведеться одночасно по декількох рядках, але умови зв'язності й обмеження ресурсів залишається єдиним для даного образу, а це значно скорочується обсяг пам'яті, необхідний для обчислень.

Висновки

Розвиток методів топологічного аналізу, кореляційних методів, методів еластичного еталона й алгоритмів нейронних мереж з використанням відомих методів теорії керування дозволяє припустити можливість створення систем розпізнавання оптичних образів, інваріантних щодо збурювань у просторі зображень.

Література

- [1] Храпливый А.П., Рожков С.А., Терновая Т.И., Бражник Д.А. Использование метода компенсации информационных потоков при создании систем автоматического распознавания//Автоматика, Автоматизация. Электротехнические комплексы и системы. - 1998, №2. С. 93-100.
- [2] Дж. Ту, Р. Гонсалес. Принципы распознавания образов. Пер. с англ. И.Б. Гуревича, под ред. Ю.И.Журавлева -М.: Мир, 1978, -с.407.
- [3] Хорн Б.К.П. Зрение роботов : Пер. с англ. - М.: Мир, 1989.-487 с., ил
- [4] Катус Г.П. Восприятие и анализ оптической информации автоматической системой. -М...: Машиностроение, 1986. -416 с., ил.
- [5] Медведев В.С., Потемкин В.Г. Нейронные сети. MATLAB 6/ Под общ. ред. к.т.н. В.Г. Потемкина. -М.: Диалог-МИФИ, 2002. -496 с.