

Л. О. ЯЩУК

## Адаптація алгоритмів розпізнавання до афінних спотворень графічних зображень у системах автоматизації сортування посилочної пошти

### Загальна характеристика афінних спотворень графічних зображень

Афінними перетвореннями графічних зображень називають такі перетворення, при яких точки переходять у точки, прямі у прямі, паралельні прямі в паралельні прямі та зберігається інцидентність точок і прямих. До афінних перетворень відносяться поворот, зсув і масштабування графічних зображень.

Афінні перетворення та відповідні спотворення графічних зображень поштових індексів становлять головну перешкоду в побудові систем автоматизованого сортування поштових відправлень.

Якщо при сортуванні легкої пошти (письмова кореспонденція) її незначна маса й стандартні розміри дають змогу достатньо чітко фіксувати положення поштових індексів відносно зчитувача, то в разі важкої пошти (посилки, пачки газет), яка має значну масу й нестандартні розміри, досягти такого фіксованого положення не вдається, а отже, постає проблема забезпечення інваріантності розпізнавання щодо зсувів, поворотів і масштабування поштових індексів у площині, перпендикулярній до оптичної осі зчитувача.

Для компенсації афінних спотворень звичайно визначають їхні параметри та виконують обернені афінні перетворення, в результаті яких графічні зображення набирають свого первісного вигляду.

Перспективною є заміна обернених афінних перетворень адаптацією алгоритмів розпізнавання до афінних перетворень цифр поштового індексу.

Основні труднощі такої адаптації пов'язані з відмінністю подання прямої лінії на неперервній і дискретній площинах (рис. 1).

Як впливає з рис. 1, при поданні на неперервній площині кількість точок, що утворюють відрізки  $ab$  і  $AB$  (континуум), однакова, тоді як на дискретній площині ця закономірність пору-

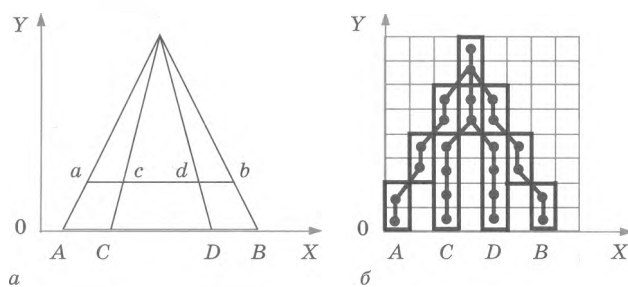


Рис. 1. Подання прямих ліній на неперервній (а) і дискретній (б) площинах

шується. Це означає, що будь-яким точкам  $C$  і  $D$  відрізка  $AB$  на неперервній площині взаємно однозначно відповідають точки  $c$  і  $d$  відрізка  $ab$ , а на дискретній площині такої взаємно однозначної відповідності немає.

Приклади зображення двох похилих ліній I і II на неперервній і дискретній площинах наведено на рис. 2, де  $\varphi_1$  і  $\varphi_2$  — кут нахилу відповідно лінії I і II ( $\varphi_1 < 45^\circ$ ;  $\varphi_2 > 45^\circ$ ).

Як бачимо, похила лінія на дискретній площині має вигляд ступінчастої лінії, в якій довжини сходинок можуть відрізнятися на одиницю.

Зображення лінії на дискретній площині надалі називатимемо її *траєю*, а кількість дискретних елементів, з яких ця траєкторія складається, — її *дискретною довжиною*.

Траєкторія відрізка прямої лінії, нахиленої під кутом  $\varphi \leq 45^\circ$ , дискретна довжина якої визначається проекцією цього відрізка на вісь абсцис, називатимемо-

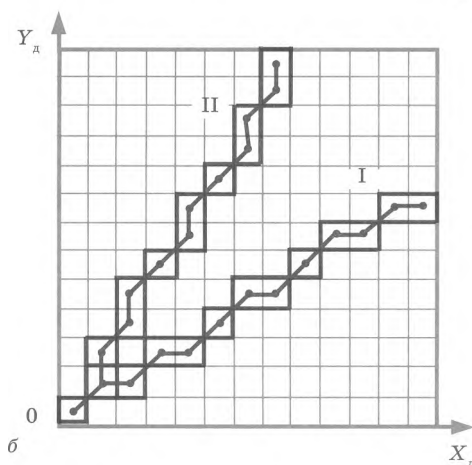
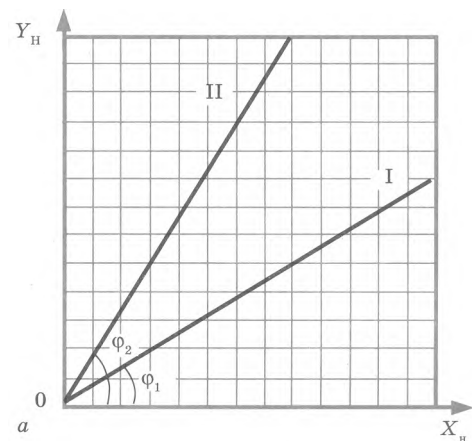


Рис. 2. Подання двох похилих ліній на неперервній (а) і дискретній (б) площинах

мо трасою типу X, а трасу відрізка прямої лінії, нахиленої під кутом  $\varphi \leq 45^\circ$ , дискретна довжина якої визначається проекцією цього відрізка на вісь ординат, — трасою типу Y.

Траса, яка на рис. 2 відповідає куту нахилу  $\varphi_1 = \text{arctg} \frac{8}{13} < 45^\circ$ , є трасою типу X, а та, що відповідає куту нахилу  $\varphi_2 = \text{arctg} \frac{13}{8} > 45^\circ$ , — трасою типу Y. Зауважимо, що оскільки  $\varphi_2 = 90^\circ - \varphi_1$ , дискретні довжини обох трас збігаються і дорівнюють 13.

Зауважимо, що дискретна довжина лінії при її повороті від 0 до  $45^\circ$  пропорційна до косинуса кута повороту, а при повороті від  $45^\circ$  до  $90^\circ$  — до синуса цього кута (рис. 3). При цьому вважають, що максимальні значення досягаються при кутах повороту 0 і  $90^\circ$ , а мінімальне — при куті повороту  $45^\circ$  (рис. 4).

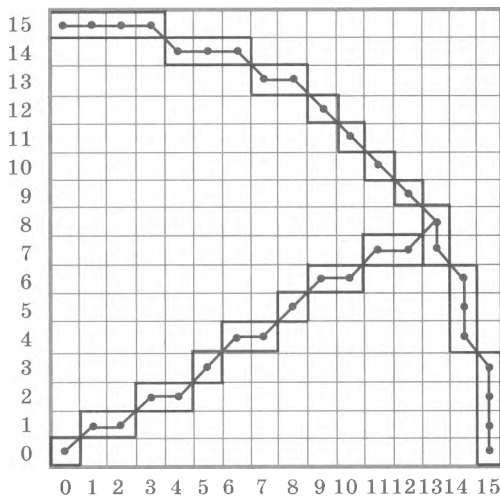


Рис. 3. Положення кінців відрізків лінії при її повороті від 0 до  $90^\circ$

У разі зображення на дискретній площині будь-якої лінії існує лише вісім напрямів переходів між двома її сусідніми елементами (рис. 5).

Основні задачі адаптації алгоритмів розпізнавання до афінних перетворень графічних зображень полягають у визначенні:

- визначення параметрів афінних спотворень;
- визначення трас переходів між елементами графічних зображень на дискретній площині.

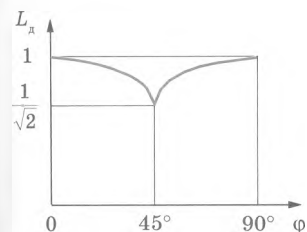


Рис. 4. Графік зміни дискретної довжини лінії залежно від кута її повороту

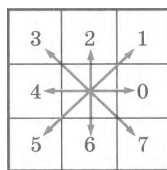


Рис. 5. Можливі напрямки переходів між сусідніми елементами лінії

### Визначення параметрів афінних спотворень графічних зображень

Параметри афінних спотворень можна визначити, проаналізувавши взаємне розміщення та відстані між спеціальними мітками (реперами), надрукованими на поштових відправленнях або на адресних ярликах прямокутної форми, що на них наклеюються.

Задача виділення міток постає також при пошуку лицьового боку поштових відправлень. Якщо за відсутності афінних спотворень (письмова кореспонденція) міток зручно використовувати прямокутні стартові та центрувальні мітки, то в разі їх наявності (посилки) як мітки доцільно використовувати геометричні фігури, інваріантні до поворотів графічних зображень.

Такими фігурами є круги або концентричні кільця. Спинимось на останніх, оскільки саме мітки у вигляді концентричних кілець (залежно від їхньої ширини і взаємної відстані) доволі просто можна ідентифікувати.

Беручи до уваги необхідність перевірки попадання адресного ярлика в зону зчитування індексу, доцільно застосувати чотири мітки, центри яких розташовані в чотирьох кутових точках ярлика (рис. 6).

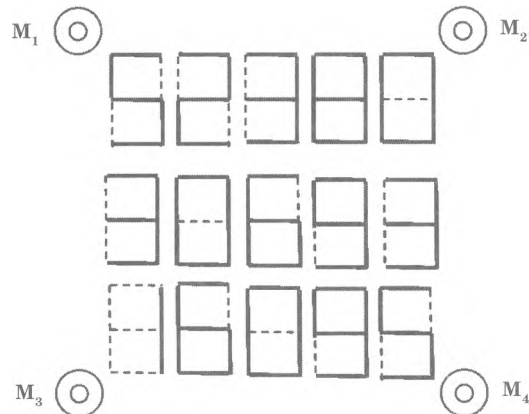


Рис. 6. Загальний вигляд адресного ярлика

У верхньому ряду ярлика записується індекс одержувача, у нижньому — індекс відправника, у середньому — номер поштового відправлення і контрольна цифра.

Значення контрольної цифри визначається як доповнення суми всіх інших цифр до числа, кратного 10.

У наведеному прикладі значення контрольної цифри (праворуч у середньому ряду) дорівнює 3, оскільки сума решти цифр дорівнює 57.

Якщо при зчитуванні адресного ярлика виявиться, що сума всіх його цифр, включаючи контрольну, кратна 10, вважається, що їх розпізнано правильно. В іншому разі якась одна або кілька цифр розпізнані неправильно, унаслідок чого відбувається відмова від розпізнавання.

За наявності однієї контрольної цифри виявляються всі окремі помилки, а також ті помилки більшої кратності, що не компенсують одна одну, наприклад:

■ вихідні цифри 38, розпізнані цифри 59; сума  $3 + 8 = 11$ , сума  $5 + 9 = 14$ ; суми не збігаються, помилки виявляються;

■ вихідні цифри 38, розпізнані цифри 56; сума  $3 + 8 = 11$ , сума  $5 + 6 = 11$ ; суми збігаються, помилки не виявляються.

Усього за допомогою однієї контрольної цифри виявляється близько 90% усіх можливих помилок.

Кожна мітка містить два концентричних кільця і проміжок між ними, які при перетині лінією діаметра утворюють послідовність семи чорних Ч і білих Б сегментів  $Ч_1$   $Б_2$   $Ч_3$   $Б_4$   $Ч_5$   $Б_6$   $Ч_7$  (рис. 7).

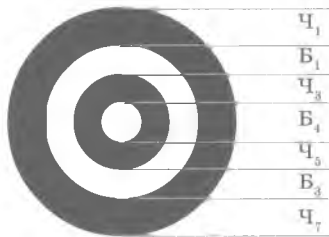


Рис. 7. Загальний вигляд мітки

Щоб можна було використати загальний алгоритм виділення міток, у них мають бути однакові діаметри, а це можливо за умови рівності сумарної кількості елементів у сегментах  $Ч_1$ ,  $Б_2$  і  $Ч_3$ .

Залежно від кількості елементів, що утворюють зазначені сегменти, вони можуть бути широкими Ш, середніми С або вузькими В, причому кількість елементів у цих сегментах задовольняє пропорцію 3 : 2 : 1.

Згідно з цим за умовні довжини сегментів візьмемо  $Ш = 3$ ,  $С = 2$ ,  $В = 1$ . Тоді для виконання рівності сумарної кількості елементів у розглянутих сегментах необхідно, щоб суми їхніх умовних довжин дорівнювали шести.

Послідовність довжин сегментів  $Ч_1Б_2Ч_3$  утворює код діаметра мітки. Усього таких кодів існує сім:

- код 1 — ВСШ — 123,
- код 2 — ВШС — 132,
- код 3 — СВШ — 213,
- код 4 — ССС — 222,
- код 5 — СШВ — 231,
- код 6 — ШВС — 312,
- код 7 — ШСВ — 321.

Ідентифікуються мітки в такий спосіб: мітці  $М_1$  присвоєно код діаметра 1, мітці  $М_2$  — код діаметра 2, мітці  $М_3$  — коди діаметрів 3, 4, 5, мітці  $М_4$  — коди діаметрів 6, 7. Завдяки цьому, окрім ідентифікації міток, за комбінацією кодів діаметрів міток  $М_3$  і  $М_4$  можна не лише ідентифікувати мітки, а й визначити шість типів поштових відправлень, наприклад: авіа, з оголошеною цінністю, з оплаченою доставкою, з післяплатою і т.ін.

Щоб переконатись у правильності розташування міток і відсутності так званих проєктивних спотворень графічного зображення (зумовлених неперпендикулярністю оптичної осі зчитувача до площини графічного зображення), за наявності яких паралельні прямі сприймаються як збіжні або розбіжні, виконують перевірку довжин проєкцій сторін адресного ярлика на осі координат (рис. 8):

$$\begin{cases} |x_1 - x_2| = |x_3 - x_4|, \\ |y_1 - y_2| = |y_3 - y_4|, \\ |x_1 - x_3| = |x_2 - x_4|, \\ |y_1 - y_3| = |y_2 - y_4|, \end{cases}$$

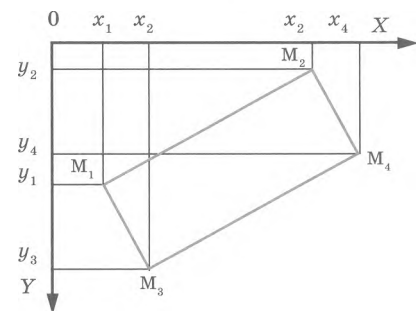


Рис. 8. Визначення довжин проєкцій сторін адресного ярлика на осі координат

а також за допомогою спеціального алгоритму визначають параметри афінних спотворень графічного зображення (рис. 9).

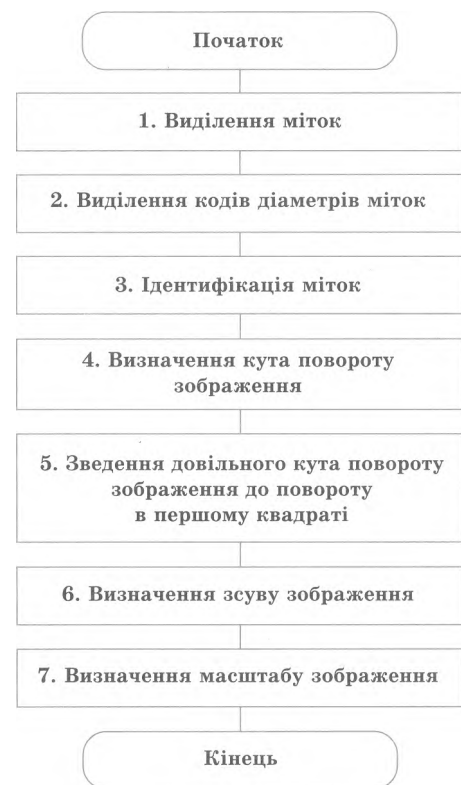


Рис. 9. Схема алгоритму визначення параметрів афінних спотворень

Фіксуючи послідовності сегментів, які відповідають вимогам щодо діаметрів міток, обчислюємо довжини  $L_1, L_2, L_3, L_4, L_5, L_6, L_7$  семи поточних сегментів кожного рядка графічного зображення, починаючи з чорного. Якщо в рядку менш як сім сегментів, то в ньому пошук міток не виконується.

За наявності в рядку семи і більше сегментів перевіряють можливість подання ними діаметрів міток ( $L_1 = L_7, L_2 = L_6, L_3 = L_5$ ). Коли зазначені умови виконуються, обчислюють сумарну довжину  $D$  усіх сегментів, яка має задовольняти нерівності:

$$D_{\min} \leq D \leq D_{\max}$$

Знайдену в такий спосіб послідовність сегментів беруть за діаметр мітки та обчислюють координати її центрального елемента.

Визначення кодів діаметрів міток зводиться до аналізу довжин сегментів  $Ч_1, Ч_2, Ч_3$ , що утворюють коди діаметрів міток.

Перевірка правильності розташування міток зводиться, як вже зазначалося, до перевірки рівностей відповідних проекцій сторін адресного ярлика на осі координат.

Параметри афінних спотворень графічного зображення визначаються за координатами центральних елементів і діаметрами міток. При цьому вважається, що поворот і зсув графічного зображення внаслідок його афінних спотворень здійснюється відносно центрального елемента мітки  $M_1$  (рис. 10).

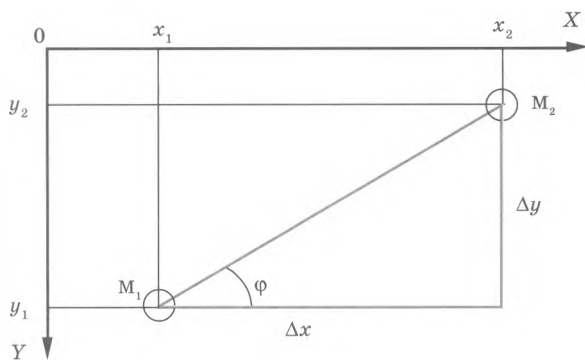


Рис. 10. Визначення кута повороту графічного зображення за взаємним розташуванням центральних елементів міток  $M_1$  і  $M_2$

З рис. 10 маємо:

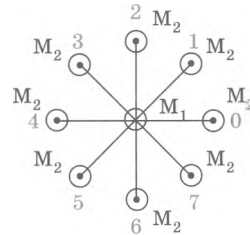
$$\varphi = \arctg \frac{\Delta y}{\Delta x}$$

На практиці кут повороту найзручніше визначати безпосередньо за приростами координат:

$$\Delta x = x_2 - x_1, \quad \Delta y = y_2 - y_1$$

Для визначення номера квадранта, в якому знайдеться відрізок між центральними елементами міток  $M_1$  і  $M_2$ , використовують знаки  $\Delta x$  і  $\Delta y$ .

Послідовність положень міток  $M_1$  і  $M_2$  в усіх квадрантах, наведено для визначення номера квадранта, наведено на рис. 11.



Номер квадранта	Приріст координати	
	$\Delta x$	$\Delta y$
1	$> 0$	$\leq 0$
2	$\leq 0$	$< 0$
3	$< 0$	$\geq 0$
4	$\geq 0$	$> 0$

Рис. 11. Послідовність положень міток  $M_1$  і  $M_2$  в усіх квадрантах

При адаптації алгоритмів розпізнавання до афінних спотворень графічних зображень зручно довільний кут повороту графічного зображення звести до кута його повороту в квадранті 1.

Із цією метою необхідно обчислити константу квадранта (КК), що (з урахуванням рис. 5) визначається так:

- квадрант 1 – КК = 0,
- квадрант 2 – КК = 2,
- квадрант 3 – КК = 4,
- квадрант 4 – КК = 6.

Напрямок  $N$  руху у квадранті 1 пов'язаний із напрямом  $N_\phi$  фактичного руху вздовж відрізка між центральними елементами міток  $M_1$  і  $M_2$  та константою квадранта КК таким співвідношенням:

$$N = N_\phi - \text{КК}$$

Наприклад:  $N_\phi = 3, \text{КК} = 2, N = 1; N_\phi = 6, \text{КК} = 6, N = 0$ .

Зсув графічного зображення визначається координатами  $x_1, y_1$  центрального елемента опорної мітки  $M_1$ .

Масштаб  $M$  графічного зображення визначається відношенням виявленого значення довжини діаметра мітки  $D_B$  до його номінального значення  $D_H$ :

$$M = \frac{D_B}{D_H} \leq 1$$

У процесі визначення параметрів афінних спотворень формується відповідь «Відмова від розпізнавання», якщо:

- кількість виявлених міток менша або більша від 4;
- довжини діаметрів виявлених міток виходять за допустимі межі;
- коди діаметри виявлених міток не дають змоги їх ідентифікувати;
- умови правильності розташування міток не виконуються.

*Урахування параметрів афінних спотворень в алгоритмах розпізнавання графічних зображень*

Будь-які алгоритми розпізнавання графічних зображень, зокрема й рукописних цифр, так чи інакше пов'язані з обходом елементів цих зображень. Оскільки в пам'яті ЕОМ подається матриця двокоординатного зображення то найчастіше як напрями обходу використовують горизонтальний (по рядках) і вертикальний (по стовпцях).

Під впливом афінних спотворень напрями обходу елементів графічного зображення та їхні дискретні довжини змінюються і мають бути відповідно скориговані.

Враховуючи, що напрями обходу елементів графічного зображення по спотворених горизонталях і по спотворених вертикалях відрізняються лише константою квадранта  $KK = 2$ , для визначення цих напрямів можна використати єдиний алгоритм.

Основна задача врахування параметрів афінних спотворень в алгоритмах розпізнавання графічних зображень полягає в побудові трас обходу елементів графічного зображення по спотворених горизонталях і по спотворених вертикалях, тобто в побудові послідовностей напрямів переходу між сусідніми елементами спотвореного графічного зображення, що відповідають напрямам обходу по горизонталях і по вертикалях первісного графічного зображення.

При цьому необхідно забезпечити щільне прилягання сусідніх трас такого обходу, тобто досягти такого взаємного розташування цих трас, при якому жодний елемент графічного зображення не буде пропущено і жодний елемент не буде враховано двічі.

Принцип побудови трас обходу елементів спотвореного графічного зображення, що задовольняють зазначені вимоги щільного прилягання, полягає ось у чому.

Згідно із зображенням нахилених ліній на дискретній площині (див. рис. 2) будують базову трасу обходу елементів графічного зображення, що сполучає центральні елементи міток  $M_1$  і  $M_2$ , і дві обмежувальні траси, що сполучають центральні елементи міток  $M_1$  і  $M_3$  та  $M_2$  і  $M_4$ .

Сусідня траса обходу елементів графічного зображення будується паралельним перенесенням усіх елементів базової траси на один елемент у вертикальному напрямі згори вниз, якщо кут  $\varphi$  нахилу лінії  $M_1M_2$  задовольняє нерівність  $\varphi \leq 45^\circ$ , або на один елемент у горизонтальному напрямі зліва направо, якщо  $\varphi \leq 45^\circ$ .

Якщо при зазначеному перенесенні крайній елемент формованої траси ходу елементів графічного зображення виходить за обмежувальну трасу, він відкидається, а якщо виявляється нестача крайнього елемента, — він доповнюється згідно з рис. 2.

Приклади побудови трас обходу елементів графічного зображення наведено на рис. 12.

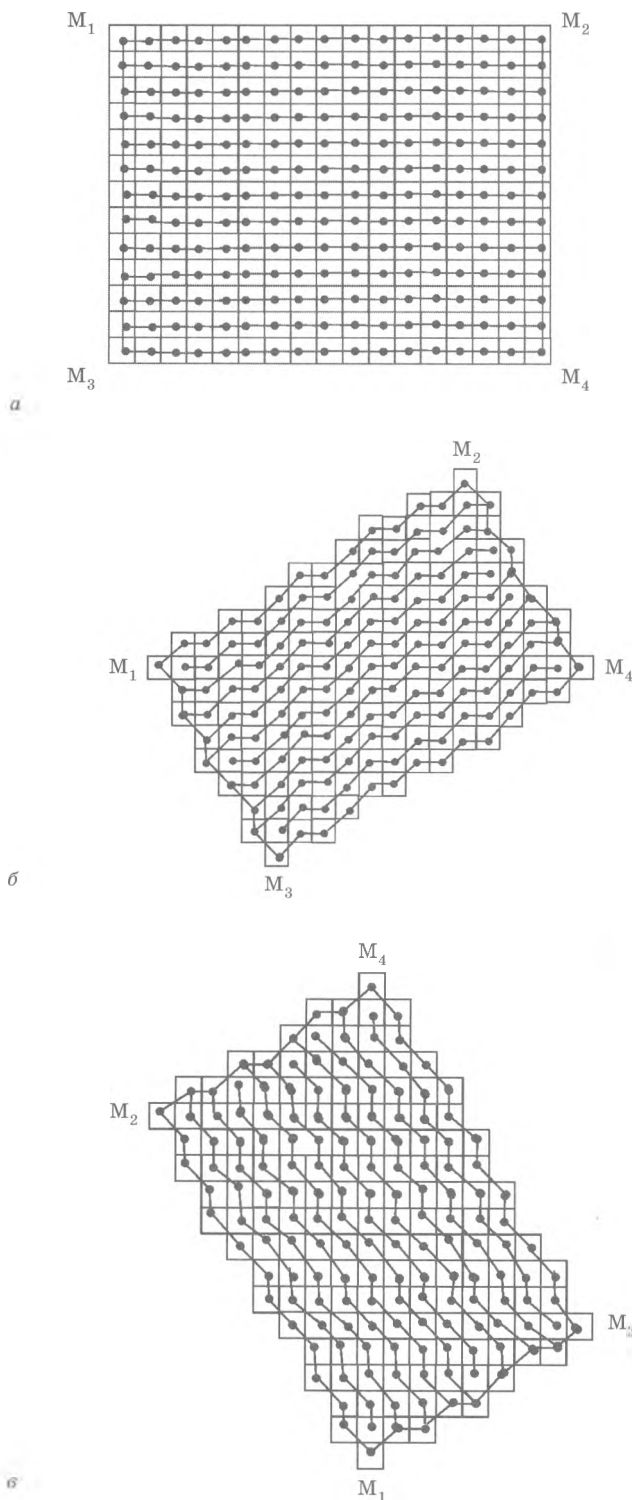


Рис. 12. Приклади побудови трас обходу елементів вихідного (а) і спотвореного графічного зображення: б — при  $\varphi \leq 45^\circ$ ; в — при  $\varphi \geq 45^\circ$

Для побудови трас обходу елементів спотвореного графічного зображення розроблено спеціальний алгоритм, що містить 11 блоків (рис. 13).

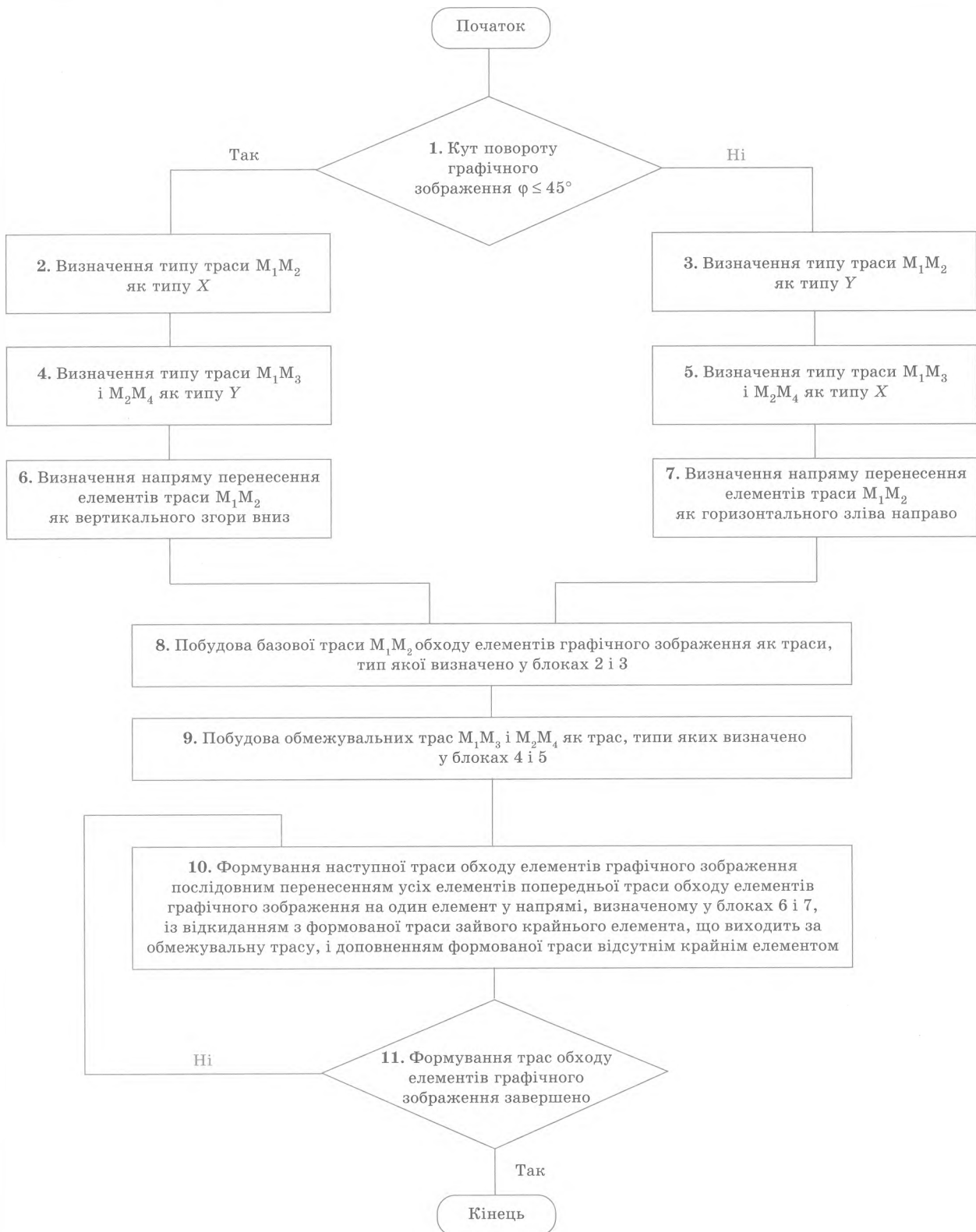


Рис. 13. Схема алгоритму побудови трас обходу елементів спотвореного графічного зображення

\* \* \*

Адаптація алгоритмів розпізнавання до афінних спотворень графічних зображень дає змогу розв'язати важливу задачу автоматизації сортування посилючої пошти, компенсації впливу зазначе-

них спотворень за рахунок використання сучасних комп'ютерних технологій, а також виявити наявність проєктивних спотворень графічних зображень поштових індексів, унеможлививши їх помилкове розпізнавання.