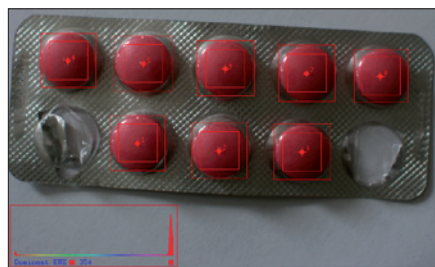


Identifikace objektů v prostředí strojového vidění VisionLab

Vizuální inspekce se stává stále běžnější součástí systémů průmyslové automatizace. S novou generací kamer, a především díky současnému programovému vybavení je sice její realizace snadnější, než tomu bylo kdykoliv dříve, ale naproti tomu se touto činností také zabývá neustále rostoucí počet lidí, pro které je problematika strojového vidění pouze okrajovým zájmem. Proto se lze v tomto oboru setkat s výrazně větším počtem neúspěšných řešení než v ostatních oblastech průmyslové automatizace.



Obr. 1. Nalezení specifických objektů na základě barev

Pro úspěšné řešení systému zakázek je důležitá volba základních principů a koncepce. I zkušenému odborníkovi může značně pomoci rezerva nejdůležitějších parametrů. Často se velmi vyplatí použít kameru s co nejvyšší kvalitou obrazu a kvalitní softwarový systém strojového vidění. Také je dobré nechat si rezervy ve výpočetním výkonu zařízení, na kterém běží algoritmy úlohy.

Systém strojového vidění VisionLab je trvale rozvíjen a postupně do něj přibývají stále náročnější a složitější algoritmy zpracování obrazu. Mnoho často používaných algoritmů lze rozumně implementovat jen díky využití mohutného masivně paralelního výkonu současných grafických procesorů. Při řešení úloh strojového vidění je nutné nepřetržitě hlídat čas, který úloha „spotřebává“. Významně zjednodušit objektivně složité algoritmy není vždy možné, a tak často nezbyvá nic jiného než použít výkonnější hardware. Vývoj v oblasti grafických procesorů si již dlouhodobě zachovává velkou dynamiku. Je sice pravda, že je poháněn průmyslem počítačových her, a nikoliv oborem strojového vidění, ale to při využívání schopnosti současných GPU vůbec nevadí. I s běžným počítačem s cenově velmi dostupným grafickým adaptérem je k dispozici systém, kterému se datovou propustností a masivně paralelním výkonem v plovoucí řádové čáře nevyrovná v podstatě žádné jiné dostupné zařízení.

A v poměru ceny a výkonu je situace ještě jednoznačnější.

Jedním z obtížně splnitelných požadavků v úlohách strojového vidění bývá spolehlivá identifikace objektů v obraze. Nalezení a roz-



Obr. 2. Identifikace obrazce maticového kódu

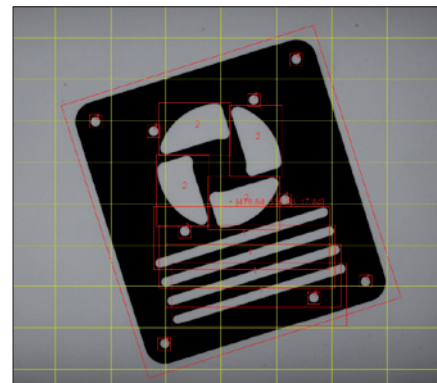


Obr. 3. Vyhledání pozice a regionu registrační značky

poznání požadovaných objektů lze považovat za podstatnou složku vedoucí k „porozumění“ obsahu obrazu strojem. Algoritmy identifikace objektů jsou často hodně náročné na spotřebu výpočetního výkonu a některé z nich

jsou rozumně použitelné jen díky využití mohutného masivně paralelního výkonu současných grafických procesorů.

Jednotlivé algoritmy se hodně liší svým zaměřením. Neexistuje sice žádný společ-



Obr. 5. Identifikace geometrických útvarů a jejich vzájemných relací

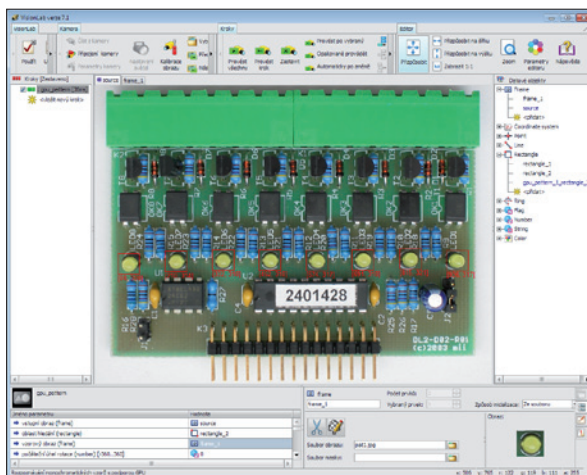
ný postup, který by řešil vše od čtení textu až po rozpoznávání obličejů, ale několik základních požadavků mají všechny postupy společné. Obvykle je nutné, aby spolehlivost rozpoznávání nebyla ovlivňována jasnou a kontrastem obrazu, často je také nutné objekty detekovat nezávisle na jejich velikosti a natočení vzhledem k souřadnému systému obrazu. Optimální algoritmy tedy musí být invariantní vzhledem k jas, měřítku i rotaci. Právě tyto požadavky a komplikace algoritmů s nimi spojené stojí za již zmiňovanou potřebou značně vysokého výpočetního výkonu.

Systém strojového vidění VisionLab pro kvalitní vyřešení těchto požadavků poskytuje množství nástrojů.

Jednoúčelové hledání specifických objektů

Sem patří mnoho často používaných metod, které nejsou opravdovou obecnou identifikací libovolných objektů. Nejjednodušší, ale hojně používaná je i prostá extrakce barevných nebo jasných skvrn či textur (obr. 1). Takto získané objekty mohou být popsány např. pomocí souřadnic svého těžiště a mohutnosti.

Za identifikaci objektů lze do určité míry považovat i např. nalezení obrazců čáro-



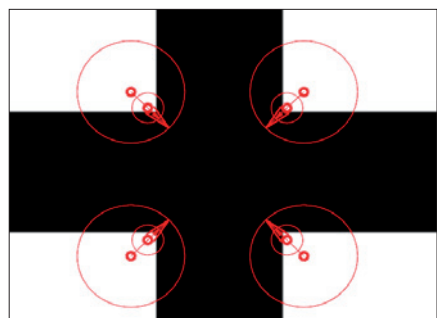
Obr. 4. Výkonná identifikace bitmapových vzorů pomocí GPU

vých a maticových (datamatrix) kódů v ploše obrazu (obr. 2). Obdobnou úlohou je i nalezení písmen a přečtení textů. Speciálně zaměřeným hledáním a identifikací písmen a číslic je i čtení registračních značek automobilů (obr. 3). Také zde je třeba v širokém rozsahu jasů, měřítek a natočení značku nalézt a poté spolehlivě přečíst všechny znaky.

Nyní ale již pojďme k obecným identifikacím (téměř) libovolných objektů.

Hledání podle bitmapových vzorů

Jde o postup v současnosti nejčastěji používaný (obr. 4). Algoritmus *pattern matching* mají k dispozici téměř všechny nástroje pro strojové vidění. Bývá implementován i v tzv. inteligentních kamerách. V obraze je pomocí normalizované křížové korelace hledán bitmapově definovaný vzor. Normalizace zde řeší invarianci vzhledem k jasů. Potíž je ale v tom, že výpočet korelačního koeficientu je natolik náročný, a to dokonce i pro nejvýkonnější současné GPU, že nemůže být dělán hrubou silou pro všechny body vzoru ani pro všechny body prohledávaného obrazu. Zde musí nastoupit nejrůznější aproximace a optimalizace, které výsledek do určité míry poškozuji. Právě tyto postupy optima-



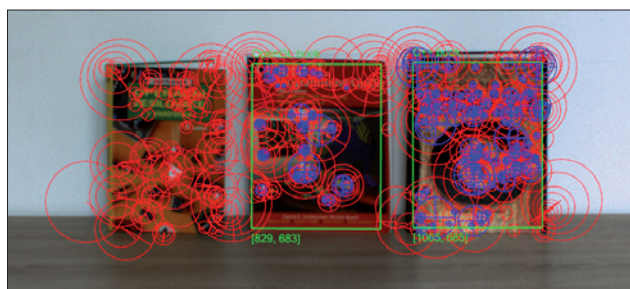
Obr. 6. Ukázka nalezených významných bodů včetně jejich orientací a měřítek na jednoduchém testovacím obrazci

lizací si dodavatelé pečlivě chrání. I systém VisionLab má ve svých víceprůchodových algoritmech množství unikátních optimalizací, které uspoří spoustu výpočetního výkonu při minimalizovaném vlivu na kvalitu a reprodukovatelnost identifikace objektů. Opravdové potíže se požadavky na výpočetní výkon nastanou u většiny implementací tehdy, bude-li požadována také invariance vzhledem k větším rozsahům měřítek a rotací.

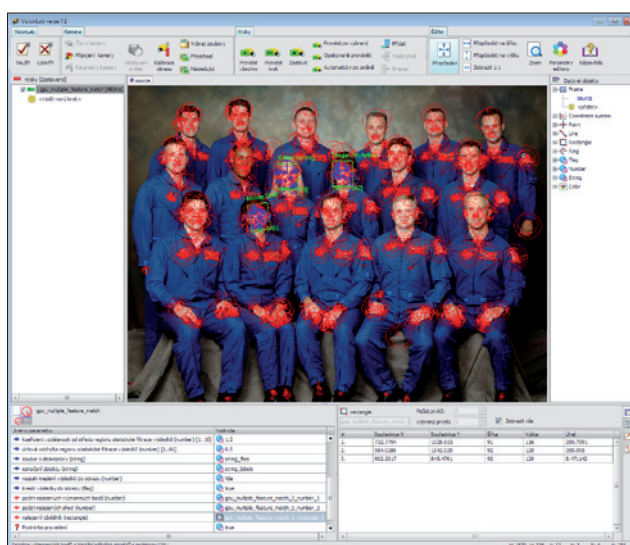
Možnosti hledání bitmapových vzorů jsou hodně široké, ale tento algoritmus také není vhodný pro všechny typy objektů. Není dobré jej používat pro černobílé objekty se strmými kontrastními hranami a problémy často způsobují i objekty s periodicky se vyskytujícími obrazci (např. nožičky integrovaných obvodů atd.). Tou hlavní potíží však zůstává náročnost na výpočetní výkon.

V systému VisionLab je implementace algoritmu hledání vzorů k dispozici v krocích *pattern_match_monochrome* a *pattern_*

match_color. Masivně paralelní řešení s podporou GPU je obsaženo v kroku *gpu_pattern*. U moderních grafických adaptérů může být tento krok i několikanásobně rychlejší než na vícejádrových CPU.



Obr. 7. Identifikace obálek knih i na nekvalitním snímku metodou deskriptorů významných bodů



Obr. 8. Nalezení a identifikace požadovaných obličejů ve skupině lidí pomocí deskriptorů



Obr. 9. Spolehlivé nalezení požadované osoby i na nekvalitně snímané fotografii

Hledání podle geometrických tvarů

Avšak i uvedená metoda rozpoznávání objektů (obr. 5) má své přednosti a nevýhody. Nejprve je nutné ze vzorového obrazu extrahovat množinu jednoduchých ge-

ometrických tvarů, které objekt dostatečně identifikují. Nelze tedy použít barevný či šedotónový obraz, přesněji řečeno, během zpracování obrazu dochází k jeho prahování. V binárním obraze je nutné nej-

prve extrahovat hrany a z nich vybrat ty, které lze považovat za obrysy hledaných geometrických útvarů. Nejčastějšími útvary jsou úsečky a kružnice včetně kruhových oblouků. Zásadní nevýhodou je omezení na binární obraz, předností je naopak přirozená invariance vzhledem k měřítku a rotaci.

Pro hledání objektů podle geometrických útvarů nabízí prostředí VisionLab krok *geometric_matching*.

Hledání podle deskriptorů v okolí významných bodů

Tyto moderní algoritmy se dostávají ke slovu teprve v posledních několika letech. Přinášejí mnoho skvělých vlastností, nicméně jsou také velmi náročné na výpočetní výkon. Rozumná implementace se neobejde bez co největšího využití výkonu GPU.

Základní myšlenkou je nezkoumat obrazové body celého obrazu, ale omezit se pouze na obrazové body (*pixel*) v okolí tzv. významných bodů. A navíc se ani v tomto okolí nepoužívá normalizovaná křížová korelace bodů obrazů, ale jsou vzájemně porovnávány vektory, tzv. deskriptory. K identifikaci objektů tedy postačí zapamatovat si pouze určitý nevelký počet těchto deskriptorů. Proces identifikace se pak již více podobá hledání podobnosti číselných řad v databázích.

Nejprve je však nutné významné body v obraze najít. Obvykle jsou to body, v jejichž okolí se mění jas v několika směrech (obr. 6). To zní celkem jednoduše, ale pod-

statnou komplikací je nutnost vyhledat významné body co nejrychleji v mnoha měřítcích obrazu. Pro každý bod je zjišťována i jeho úhlová orientace – vektor největšího gradientu jasu v jeho okolí úměrném měřítku. Tato informace zajišťuje invarianci vzhledem k rotaci. Pro oblasti kolem detekovaných významných bodů se již jen vypočítají vektory deskriptorů. Deskriptory charakterizují směry a dynamiku gradientů jasu obrazové funkce kolem významných bodů.

Při vyhledávání objektů pak již jen postačí identifikovat zapamatované deskriptory s nově nalezenými.

V systému strojového vidění VisionLab jsou pro využití těchto algoritmů k dispozici kroky *gpu_keypoint*, *gpu_feature_save*, *gpu_feature_match* a *gpu_multiple_feature_match*. Možnosti těchto kroků jsou překvapující. Kromě invariance vzhledem k jasu, měřítku a neomezené rotaci umožňují s vysokým výkonem vyhledávat a identifikovat v podstatě veškeré typy a velikosti objektů. Metoda je velmi odolná i proti obrazovému šumu, změnám ostrosti obrazu a částečnému skrytí objektů (obr. 7). Jejím prostřednictvím je možné s velkou spolehlivostí identifikovat např. i lidské obličeje apod. (obr. 8, obr. 9).

Společná softwarová platforma pro strojové vidění i průmyslovou automatizaci

Díky využití čistých a stabilních surových obrazových dat z kamer DataCam, jednotné softwarové platformě Control Web a systému VisionLab s implementací pokročilých algoritmů využívajících vícejádrové CPU i masivně paralelní GPU mají autoři zakázkových aplikací k dispozici systém, který významně zvyšuje pravděpodobnost úspěšného řešení i u velmi složitých úloh.

Roman Cagaš, Moravské přístroje, a. s.