

LitterDrone: monitorización de basura marina empleando drones y análisis de imagen

Fernando Martín-Rodríguez⁽¹⁾, Laura Rodríguez-Barreiro⁽¹⁾, Marta Fernández-Bastos⁽¹⁾,
Gema Martínez Iglesias⁽¹⁾.

E-mail: fmartin@tsc.uvigo.es, E-mail segundo autor, E-mail tercer autor, etc.

⁽¹⁾ Dpto. de Teoría de la Señal y Comunicaciones. Universidad de Vigo.

E.E.T. C/Maxwell S/N (Ciudad Universitaria). 36310 Vigo (Pontevedra).

Abstract- This communication is about “LitterDrone” project. LitterDrone is funded via the Blue-Labs program of the European commission and it aims to make a contribution to solve the problem of marine litter. Part of this problem is monitoring stranded marine litter on beaches (measuring number and type of litter elements). Monitoring results can be used to infer data on litter origin and on the influence of tides, currents and human activity. OSPAR convention [1] is a joint European initiative that tries to unify forces against marine pollution. Part of this convention implies that contracting parties (countries) must monitor periodically stranded marine litter on beaches. Spain has signed the convention in January 1994. Litter monitoring in Spain is nowadays implemented by human personnel counting (& picking) litter items in certain beaches at certain times (4 campaigns each year, one for each season). LitterDrone project aims to create a new and/or complementary methodology based on obtaining images from drone flights (creating orthomosaics of RGB and multispectral images) and developing software to analyze such images to obtain results comparable to those of the manual sampling.

I. INTRODUCCIÓN

LitterDrone tiene como objetivo el desarrollo de una herramienta innovadora para el control y gestión de la basura marina, aplicando para ello nuevas tecnologías a la monitorización ambiental, como los vehículos aéreos no tripulados (popularmente conocidos como “drones”) equipados con cámaras de alta resolución, y software especializado para el procesado de las imágenes obtenidas por los drones, de cara a la detección y caracterización de los residuos detectados en las playas.

El proyecto LitterDrone está siendo desarrollado por un consorcio constituido por la Universidad de Vigo, coordinadora del proyecto, la empresa Grafinta S.A. y la Asociación Española de Basuras Marinas (AEBAM). El proyecto, cuenta con el apoyo de ECOEMBES y con la colaboración del Parque Nacional Marítimo-Terrestre de las Islas Atlánticas de Galicia (PNIAG).

Para el trabajo de campo del proyecto (vuelos) se han seleccionado dos playas concretas, ambas situadas en la isla central del archipiélago de las Cíes. Se trata de:

- La playa de Rodas que es una de las playas monitorizadas periódicamente por el MAPAMA (Ministerio de Agricultura, Pesca, Alimentación y Medio Ambiente) a través de muestreos a pie.
- La playa de “Nosa Señora” que es una playa de menor dimensión cercana a la anterior.

En ambas se han realizado (y quedan por realizar) campañas de vuelos que consisten en la obtención de imágenes aéreas. A veces, con residuos colocados deliberadamente, otras con los residuos presentes inicialmente. Los resultados de la nueva metodología se comparan con los obtenidos mediante muestreos a pie.

II. OBTENCIÓN DE IMÁGENES

Las imágenes se obtienen con una plataforma tipo UAV volando con auto-piloto sobre el tramo de interés. Se han probado diferentes combinaciones plataforma-cámara siendo las más útiles:

- Plataforma: multirrotor volando entre 10 y 20 m de altura. Por encima de esta altura la resolución se hace menor a 1 cm por píxel y no es posible la detección de residuos de pequeño tamaño. Se probaron plataformas de ala fija con menor éxito debido a que deben volar a mayor altura. Eso sí, pueden ser útiles para detectar residuos grandes en tramos de gran longitud.
- Cámara RGB: se trata realmente de una cámara fotográfica convencional (se ha usado el modelo SONY ILCE QX1, por ser fácilmente adaptable a la plataforma UAV. Las imágenes RGB se han revelado hasta ahora como las más útiles en el proceso de detección y clasificación. Destaca por su resolución y además no es necesario alinear las bandas obtenidas (tres bandas: R, G y B).
- Cámara multispectral: se ha utilizado una cámara de 5 bandas, modelo MicaSense RedEdge. Dispone de las bandas: R, G, B, NIR (infrarrojo cercano) y RE (RedEdge: frontera entre rojo e infrarrojo). Estudiando la reflectancia de los diferentes materiales con ayuda de la biblioteca espectral USGS [2], descubrimos que por debajo de 950 nm no es posible distinguir el material. Lamentablemente, las cámaras capaces de obtener información en esas bandas son de un costo inabordable para nuestro proyecto. Por ello, se concluye que la información de infrarrojo cercano no es muy significativa. Aunque la información de la banda NIR puede usarse para detectar (y descartar como residuo) las posibles masas vegetales presentes en la playa. (ver sección III.B).

Dado que la plataforma UAV captura imágenes con un disparo periódico (una imagen cada t_0 segundos), después de

valorar varias opciones se concluyó que lo más razonable sería utilizar un programa de fotogrametría para lograr la integración de todas ellas en una única imagen global por cámara y vuelo (ortofoto u ortomosaico). En el caso de la cámara multispectral, cada disparo produce 5 fotos (una por banda), que deben ser alineadas. Hasta ahora, se ha empleado el software comercial Photomodeler [3], que alinea las componentes y crea un ortomosaico georreferenciado. Además de Photomodeler en el mundo de la fotogrametría existen otras opciones, incluidas algunas open-source.



Fig. 1. Izquierda: zonas de interés en la playa de Rodas (coinciden con los dos tramos de estudio del MAPAMA, uno corto de 100 m y otro largo de 1000 m) y en Nosa Señora se define otro tramo de 100 m. Derecha: informe gráfico de un vuelo real donde se representan todos los puntos de disparo realizados.

III. PROCESADO DE IMÁGENES

Una vez obtenidas las ortofotos, se continúa con su procesamiento para obtener un informe detallado de los residuos presentes en la zona de estudio.

Se ha desarrollado una aplicación gráfica en MATLAB para realizar esta función. Además, gracias al “MATLAB Compiler”, es posible trabajar con la aplicación sin disponer de licencia de MATLAB. El resultado final es un fichero con formato CSV (formato texto con separadores) que consiste en un listado de residuos encontrados, su posición y su clasificación. Se utilizan los códigos de categoría (material) y subcategoría (tipo de residuo) definidos por OSPAR [1]. Este informe se obtiene directamente en forma digital y sustituye al informe manuscrito generado en los muestreos a pie, que aún debe ser transcrito manualmente.

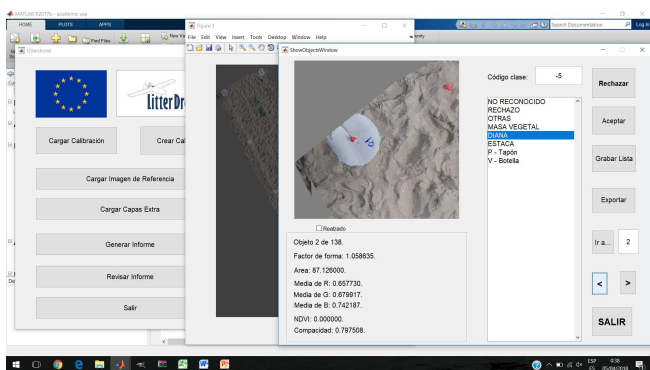


Fig. 2. Aplicación LitterDrone.

A. Localización de Objetos

Nuestra aplicación utiliza la colorimetría [4] para localizar los objetos presentes en las ortofotos. Básicamente, se trata de definir cuáles son los patrones (colores) de fondo a

eliminar. El color de un píxel se define con 6 componentes numéricas (vector de características de tamaño 6), a saber:

- Cb y Cr del espacio de color YCbCr.
- H y S del espacio de color HSV.
- a y b del espacio de color Lab.

Se trata de centrarse en las características que definen el color, eliminando la influencia del brillo.

Realizando una calibración basada en pequeñas imágenes que corresponden con los fondos a eliminar: arena seca, arena mojada, roca, agua... se caracterizan dichos fondos utilizando el conocido algoritmo K-means [5]. Reteniendo 5 prototipos (medias) por clase de fondo se utiliza la regla del vecino más próximo (1-NN) para determinar si los puntos de la imagen pertenecen a alguna de ellas o no. Dado un píxel, si el vecino más próximo dista del punto actual menos que un umbral empírico, se le considerará fondo. Todo lo que no es fondo, se considera residuo.

Sólo con este proceso ya se puede obtener un informe preliminar sobre número y densidad de residuos (figura 4).

B. Fusión de Ortofotos

En caso de tener varias imágenes del mismo tramo (por ejemplo una imagen RGB y una imagen multispectral), nos interesa hacerlas coincidir para poder utilizar la información de todas. Dado el proceso explicado en la sección anterior, siempre tomamos una imagen RGB como referencia, que además servirá para localizar los objetos. Pero, de existir otras componentes, se superpondrán a éstas para poder tener información adicional útil en la fase de clasificación. Por ejemplo, teniendo información del infrarrojo cercano (NIR) se puede calcular el índice NDVI [6]. Dicho índice permite detectar las masas vegetales (que no deben confundirse con residuos). El índice se calcula con la fórmula:

$$NDVI = \frac{NIR - R}{NIR + R}$$

Que arroja siempre un resultado entre -1 y +1, cercano a +1 cuando se trata de masas vegetales verdes, más cercano a 0 cuando se trata de masas vegetales amarillentas y muy baja al rojo es efecto de la clorofila.

La fusión es realmente un proceso de registro de imágenes donde se utiliza la foto RGB como referencia. Las componentes extra se hacen coincidir exactamente con ésta mediante transformaciones geométricas: cambio de resolución, traslación y/o giro (transformación afín de las coordenadas). La técnica de registro utilizada está basada en puntos de control (pares de puntos coincidentes en ambas imágenes). En la figura 5, se registra una imagen multispectral con una RGB. Se modifica la multispectral, hasta hacerla coincidir con la RGB (referencia). Se han usado los centros de las dianas como puntos de control. Con cuatro puntos fue suficiente (las dianas son círculos blancos colocados deliberadamente, normalmente en las esquinas de la imagen, son fijados con piquetas que se usan como posición de referencia para la georreferenciación de la ortofoto).

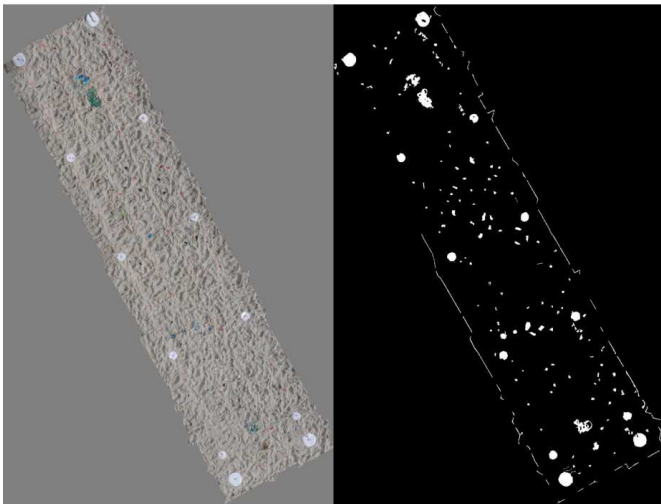
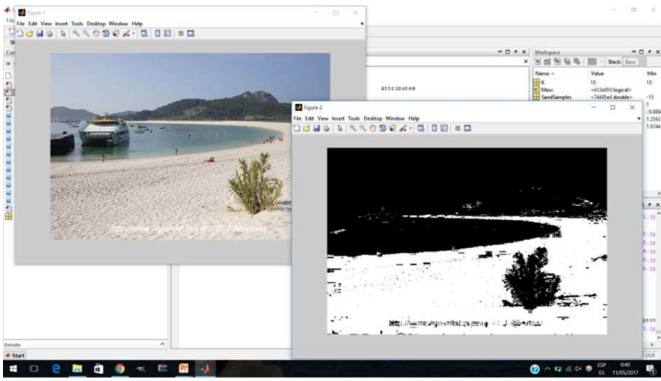


Fig. 3. Arriba: prueba preliminar del método de eliminación de fondo sobre una imagen de la playa de Rodas, arena resaltada en blanco. Centro: prueba sobre ortofoto de una cuadrícula de 17x3 m, dividida por cordeles en secciones de 1 m² (residuos colocados deliberadamente). Abajo: prueba realizada en un pabellón deportivo, se elimina el color del suelo y el de las rayas de la pista.

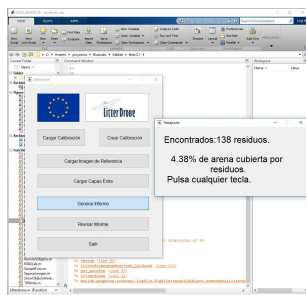


Fig. 4. Informe preliminar, generado a partir de la información de localización.

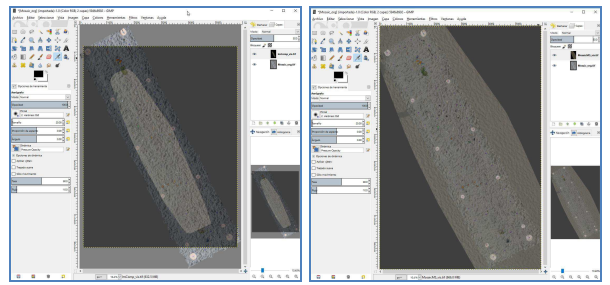


Fig. 5. Registro de una imagen multispectral sobre una RGB (referencia). Se utiliza el programa libre GIMP para poner de manifiesto que sin registro no hay coincidencia (izquierda), mientras que tras modificar la imagen multispectral se consigue su coincidencia con la base RGB (derecha). Se han usado los centros de las dianas como puntos de control, con cuatro puntos fue suficiente. Teóricamente es posible desde 3 puntos dado el número de incógnitas de la transformación afín. Al estar las dos imágenes georreferenciadas la transformación se limita a corregir desplazamiento y escala, no hay giro.

C. Clasificación de Objetos

Ésta es la parte más compleja del proyecto, ya que el convenio OSPAR define muchos tipos y subtipos de residuos. Aplicando solamente una lista simplificada, quedaríamos en: 60 clases a reconocer (un número excesivo para cualquier sistema automático).

Además, dado que con la tecnología disponible no podemos distinguir los materiales, hemos tenido que plantear un reconocimiento por imagen donde, además de un número muy alto de clases, tendremos otros problemas: objetos deteriorados, rotos, deformados y/o poco visibles (semi-enterrados). Por si fuera poco, algunos de los tipos definidos presentan una enorme variabilidad en forma y color, por ejemplos tenemos: “recipiente de bebida de vidrio” o “recipiente de bebida de plástico” presentan una variabilidad enorme.

Por eso, hemos planteado un sistema que realiza un “pre-reconocimiento” automático y presenta al usuario los objetos encontrados junto al resultado del mismo. Con pre-reconocimiento entendemos “intentar distinguir un objeto conocido”. El resultado que puede ser “rechazo” (no se reconoce nada), puede ser correcto o puede ser necesario que el usuario lo corrija. A través de una interfaz gráfica (figura 2), el usuario puede realizar la corrección de todo el informe. Básicamente, se trata de poder realizar el informe en la oficina, con la ayuda de esta herramienta, en la que ya se han reconocido correctamente ciertos ítems que sólo habrá que confirmar.

Para realizar el reconocimiento se ha implementado un banco de reconocedores simples que trabajan en paralelo intentando identificar patrones conocidos a través de características de forma y color: relación de aspecto, área,

compacidad ($c = \frac{4\pi \text{Área}}{\text{Perímetro}^2}$), distancia entre “vectores de forma” (proyección horizontal), y componentes colorimétricas (H_{med} , S_{med}). También se incluye aquí (si está disponible), el índice NDVI.

Dado un patrón de clase desconocida (imagen de entrada) se calcula una función discriminante para cada clase contemplada: D_i . Inicialmente, todas las D_i 's valen 1. Estos valores son actualizados con cada una de las características numéricas contempladas, multiplicando por una función $d_i(x)$, esto es:

$$D_i^{n+1} = D_i^n \cdot d_i(x)$$

Donde si, por ejemplo, x es la relación de aspecto (calculada para que sea mayor que 1), e i es la clase "tapón de plástico"; sabremos si es de verdad un tapón, x debe ser cercano a 1 (ligeramente mayor).

$d_i(x)$ siempre se define así:

$$d_i(x) = \begin{cases} 0, x > x_{max} \cup x < x_{min} \\ \exp \left[-1/2 \left(\frac{x - x_{med}}{x_{desv}} \right)^2 \right] \end{cases}$$

Siendo ésta una fórmula empírica pero inspirada en la distribución gaussiana y en la regla de Bayes. Para el caso del tapón y la relación de aspecto podríamos tener: $x_{min}=1.00$, $x_{max}=1.50$, $x_{med}=1.00$ y $x_{desv}=0.50$.

Nótese que puede haber objetos, como por ejemplo las masas vegetales, donde el parámetro definitorio es el índice NDVI y la relación de aspecto es irrelevante. En este caso podríamos definir: $x_{min}=0.70$, $x_{max}=1.00$, $x_{med}=1.00$ y $x_{desv}=0.25$ para el índice NDVI y para la relación de aspecto podemos definir: $x_{min}=1.00$, $x_{max}=\infty$, $x_{med}=1.00$ y $x_{desv}=\infty$, que en la práctica llevan a $d_i(x)=1$.

Esta fórmula también se usa con otras características como la forma concreta de un objeto, calculando un vector de forma a través de, por ejemplo, la proyección horizontal. En este caso, se define x (característica escalar de forma) como la distancia euclídea entre el vector de forma y la forma esperada para la clase. En este caso, la fórmula de $d_i(x)$ se modifica ligeramente:

$$d_i(x) = \begin{cases} 0, x > x_{max} \cup x < x_{min} \\ \exp \left[-1/2 \left(\frac{x}{x_{desv}} \right)^2 \right] \end{cases}$$

Para los objetos donde la forma no es relevante, no se almacena una forma esperada y se considera x igual a cero.

Mención especial merecen los objetos extremadamente delgados, como bastoncillos, hilos... Para este caso, introducimos una característica x definida como: "radio del mínimo elemento estructurante circular tal que hace desaparecer el objeto tras una operación morfológica de apertura". Dicho valor se calcula a través de sucesivas aperturas [7].

Una vez terminado el proceso, el número D_i máximo define la clase ganadora (reconocida). Por supuesto, es posible (y probable) que todos los D_i 's valgan cero y no tengamos resultado concluyente.

Aunque este método no es más que un árbol de decisión heurístico ligeramente formalizado, está resultando útil.

En imágenes que presentan todos los tipos de residuos (como la de la figura 3 donde han sido distribuidos de forma programada) se han logrado reconocer exitosamente: las dianas de georreferenciación, las estacas utilizadas para

delimitar la cuadrícula, botellas de vidrio, botellas de plástico, latas de cerveza alargadas, tapones plásticos, cordeles, bastoncillos... Se ha estudiado también cómo eliminar falsos positivos procedentes del proceso de extracción (debidos casi siempre a variaciones de color: sombras, pisadas...). Realmente, las dianas y estacas no forman parte de la clasificación OSPAR. Se han añadido con un código negativo en la lista de objetos reconocidos, de manera que se eliminan fácilmente de la lista final.

Los resultados son muy precisos en las clases contempladas (el 100% en todas excepto algunos fallos en las botellas de vidrio y plástico que se reconocen por la forma).

IV. CONCLUSIONES

LitterDrone es un proyecto que intenta mejorar los procesos de monitorización y control de la basura marina en las playas europeas. Está diseñado conociendo los principios del protocolo OSPAR [1], que es el que siguen los países de la UE, incluida España. El proyecto consiste en el desarrollo de nuevas metodologías basadas en la captura de imágenes aéreas mediante plataformas no tripuladas (drones) y diferentes tipos de cámara.

La captura se realiza utilizando tecnología conocida basada en software de fotogrametría.

El análisis de las imágenes se realiza a través de software de procesado de imagen desarrollado en entorno MATLAB.

Los métodos son sencillos y basados en técnicas de visión artificial. Sin embargo, los resultados son útiles y originales ya que no existe, hoy día, otro software capaz de un proceso similar.

LitterDrone es un proyecto a dos años que termina el 31 de enero de 2019. En el tiempo restante se planean fundamentalmente dos líneas de actuación:

- La prueba con otros tipos de sensor, como cámaras hiper-espectrales y cámaras térmicas, buscando nuevas posibilidades de discriminación de materiales.
- Mejora del reconocimiento por imagen, añadiendo nuevos patrones al sistema actual y/o probando otros métodos. No se descarta el uso de técnicas más avanzadas como las redes neuronales convolucionales, aunque de momento la base de datos existente no es suficiente para este tipo de entrenamiento.

AGRADECIMIENTOS

Los autores desean agradecer la financiación recibida de la comisión Europea a través de la acción "EASME/EMFF/2016/1.2.1.4". Además, agradecemos la colaboración financiera de Ecoembes y la ayuda prestada por el Parque Nacional de las Islas Atlánticas de Galicia.

REFERENCIAS

- [1] Convenio OSPAR: <https://www.ospar.org/>.
- [2] USGS Spectroscopy Laboratory: <https://speclab.cr.usgs.gov/spectral-lib.html>.
- [3] Photomodeler Photogrametry software: <http://www.photomodeler.com/index.html>.
- [4] Günter Wyszecki, Color science : concepts and methods, quantitative data, and formulae, New York: John Wiley & Sons, 2000.
- [5] J.A. Hartigan, Clustering algorithms, John Wiley & Sons, 1975.

- [6] James Verdin et al, Índice Diferencial de Vegetación Normalizado (NDVI), FEWS - Red de Alerta Temprana Contra la Inseguridad Alimentaria, Centroamérica, USGS/EROS Data Center, 2003.
- [7] R.C. González, Digital image processing, Upper Saddle River: Prentice Hall, cop. 2008.